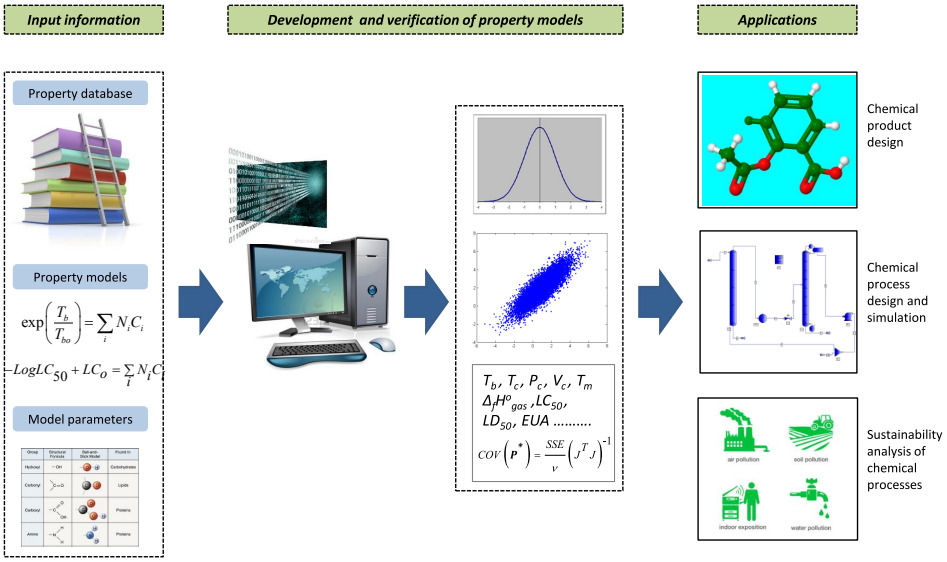
Development of pure component property models for chemical product-process design and analysis

Hukkerikar, Amol Shivajirao

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Development of pure component property models for chemical product-process design and analysis



Amol Shivajirao Hukkerikar
Ph.D. Thesis
September 2013

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Development of pure component property models for chemical product-process design and analysis

Ph.D. Thesis

Amol Shivajirao Hukkerikar

10 September 2013

Computer Aided Process-Product Engineering Center
Department of Chemical and Biochemical Engineering
Technical University of Denmark

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Address: Computer Aided Process Engineering Center
Department of Chemical and Biochemical Engineering
Technical University of Denmark
Building 229
DK-2800 Kgs. Lyngby
Denmark
Phone: +45 4525 2800
Fax: +45 4588 4588
Web: www.capec.kt.dtu.dk

Print: J&R Frydenberg A/S
København
Oktober 2013

ISBN: 978-87-93054-11-0

Preface

This thesis is submitted as a partial fulfillment of the requirements for the degree of Doctor of Philosophy (Ph.D.) in Chemical Engineering at the Technical University of Denmark (DTU). This project is a collaboration between the Computer Aided Process-Product Engineering Center (CAPEC) of the Department of Chemical Engineering, DTU and the Alfa Laval Copenhagen A/S, Denmark. The project has been carried out from July 2010 until June 2013 under the supervision of Associate Professor Gürkan Sin, Associate Professor Jens Abildskov, Dr. Bent Sarup, and Professor Rafiqul Gani.

First and foremost, I am extremely grateful to my supervisors, Associate Professor Gürkan Sin, Associate Professor Jens Abildskov, Dr. Bent Sarup, and Professor Rafiqul Gani for their guidance and scholarly inputs throughout this project. Their valuable suggestions and constructive feedback have helped me a lot during all phases of this project. A special thanks to my main supervisor Associate Professor Gürkan Sin for his continuous encouragement and support. I am highly indebted to Professor Rafiqul Gani for his guidance, motivation, and for providing numerous opportunities to collaborate with several organizations to work on interesting and challenging projects.

The thesis would not have come to a successful completion without the help I received from wonderful colleagues at CAPEC and at Alfa Laval. To all my colleagues at CAPEC, thank you. To all my colleagues at Alfa Laval, especially Dr. Bent Sarup, Alexey Shevchenko, Dr. Lars Harild, and Dr. Ling Hua, thank you for your kind support during my stay at Alfa Laval. A special thanks to Dr. Bent Sarup for providing me an opportunity to work on a challenging industrial project and for his valuable constructive feedback throughout this project. I would like to thank Dr. Rob Meier (DSM, The Netherland), Dr. Antoon Ten Kate (Akzo Nobel, The Netherland), and Dr. D. M. Young (US EPA) for their valuable inputs and for leading our collaboration to fruition. I am grateful to evaluation committee members: Dr. Michael Frenkel (NIST, USA), Dr. Gerard Krooshof (DSM, The Netherland), and Prof. Georgios Kontogeorgis (DTU Chemical Engineering) for providing feedback and valuable comments on my thesis. Special thank goes to Thomas Bisgaard for helping me in translating the thesis-summary into Danish. The work on this thesis was supported by the MULTIMOD ITN network under the European Union Seventh Framework Programme (FP-7). My dearest thanks to the people in this unique ITN network.

I am deeply and forever indebted to my parents, Shri. Shivajirao Hukkerikar and Sou. Kalawati Hukkerikar, for their love, support, encouragement at every stage of my personal and professional life. To my parents, thank you. I also thank my entire family and friends in

India for their good wishes. I am extremely grateful to my wife Anita for her unconditional love, patience, and support which has served as a secure anchor during the easy and hard times. Last but not the least, I am thankful to my sweet little daughter, Gargi, for offering a wonderful parenthood while I was working on this thesis.

Kgs. Lyngby, September 2013 Amol Shivajirao Hukkerikar

Abstract

Property prediction models based on the group-contribution$^+$ (GC^+) approach have been developed to provide reliable predictions of pure component properties together with uncertainties of predicted property values which is much needed information in performing chemical product and process design and analysis of sustainable chemical processes. For developing property models, a systematic methodology for property modeling and uncertainty analysis is employed. The methodology includes a parameter estimation step to determine parameters of the property model and an uncertainty analysis step to establish statistical information about the quality of parameter estimation, such as the parameter covariance, the standard errors in predicted properties, and the confidence intervals. For parameter estimation, large data sets of experimentally measured property values of a wide range of pure components taken from the CAPEC database, the US Environmental Protection Agency (EPA) database, and the USEtox database are used. In total, 21 thermo-physical properties and 22 environmental-related properties of pure components which include normal boiling point, critical constants, standard enthalpy of formation, liquid viscosity, fathead minnow 96-h LC_{50}, oral rat LD_{50}, global warming potential, emission to urban air (carcinogenic and noncarcinogenic) among others are modeled and analyzed. For all the estimated pure component properties, the corresponding 95% confidence intervals are also reported thereby providing information on the degree of accuracy of the property estimates. In addition, a method based on the 'molecular structural similarity criteria' is developed so that efficient use of knowledge of properties could be made in the development/improvement of property models. This method, in principle, can be applied to a wide range of properties of pure components. In this work, however, the application of 'molecular structural similarity criteria' is illustrated by considering performance improvement of models for enthalpy of formation, enthalpy of fusion, and critical temperature. For all properties listed above, it has been possible to achieve significant improvements in the performance of their models. The improved GC model for enthalpy of formation yields an average absolute deviation of 1.75 kJ/mol, which is well within the required *chemical accuracy*. Important issues related to property modeling such as: (i) quantity of property data used for the parameter regression; (ii) selection of the most appropriate form of the property model function; and (iii) the accuracy and thermodynamic consistency of predicted property values are also discussed. The developed models have been implemented into ProPred®, a property estimation toolbox of Integrated Computer Aided System, ICAS®, developed at CAPEC, DTU. Finally, a methodology for performing sensitivity analysis of process design due to uncertainties of property estimates is presented. This methodology allows the user to evaluate the effects of uncertainties of property estimates on the final design; list and rank properties that are most important from process design point of view; and establish acceptable levels of accuracy for

property models. The application of this methodology is highlighted through three case studies namely, design of an extractive distillation process, design of a short-path evaporator, and design of a de-acidification system of vegetable oil deodorization process.

Resume på Dansk

Egenskabsmodeller, baseret på group-contribution^{+} (GC^{+}) tilgangen er blevet udviklet, for at opnå pålidelige estimater af renkomponentsegenskaber samt disses tilhørende usikkerheder, hvilket er værdifulde oplysninger i forbindelse med udførelse af kemisk produkt- og procesdesign samt analyser af bæredygtige kemiske processer. En systematisk metoder er anvendt til at udvikle egenskabsmodeller og usikkerhedsanalyser. Denne metode omfatter et trin, bestående af estimering af parametre hørende til egenskabsmodellen, samt et trin, som omfatter usikkerhedsanalysen. I sidstnævnte trin etableres statistiske oplysninger om kvaliteten af parameterestimering, såsom parameterkovariansen, standardfejlen for de forudsagte egenskaber samt konfidensintervaller. Parameterestimering er baseret på eksperimentelt bestemte egenskaber af adskillige rene komponenter taget fra CAPEC database, US Environmental Protection Agency (EPA) database, og USEtox database. I alt modelleres og analyseres 21 thermo-fysiske egenskaber og 22 miljørelaterede renkomponentsegenskaber, der blandt andet omfatter normalkogepunkt, kritiske konstanter, standard dannelsesenthalpi, væskefaseviskositet, fathead minnow 96-h LC_{50}, oral rat LD_{50}, global warming potential, emission to urban air (carcinogen og ikke-carcinogen). For alle de estimerede renkomponentsegenskaber er de tilsvarende 95%-konfidensintervaller rapporteret, for at give oplysninger om estimaternes nøjagtigheder. Desuden er en metode baseret på egenskabs-data-modelanalyse blevet udviklet, således effektiv udnyttelse af viden relateret til egenskaber kan inddrages i udviklingen/forbedringen af egenskabsmodeller. I princippet kan egenskabs-data-modelanalysen anvendes til adskillige renkomponentsegenskaber, men i dette arbejde er anvendelsen af analysemetoden illustreret ved at undersøge forbedringer af modeller til estimering af dannelsesenthalpi, smelteenthalpi og kritisk temperatur. For alle de ovenfor nævnte egenskaber, har det været muligt at opnå væsentlige forbedringer af deres respektive modeller. Den forbedrede GC-model til estimering af dannelsesenthalpi giver en gennemsnitlig, absolut afvigelse på 1,75 kJ/mol, hvilket kan betragtes som godt inden for den krævede *kemiske nøjagtighed*. Følgende vigtige spørgsmål i forbindelse med egenskabsmodellering bliver også diskuteret: (i) mængde af egenskabsdata, der anvendes i parameterregressionen, (ii) udvælgelse af det mest hensigtsmæssige matematiske udtryk til egenskabsmodellen, og (iii) rigtigheden og den termodynamiske konsistens af estimerede egenskaber. De udviklede modeller er blevet inkorporeret i ProPred®, en egenskabsestimeringsværktøjskasse i Integrated Computer Aided System, ICAS®, udviklet af CAPEC, DTU. Endelig, er en metode til at udføre følsomhedsanalyse af procesdesign på grund af usikkerheder af egenskabsestimater præsenteret. Denne metode gør det muligt for brugeren at vurdere virkningerne af usikkerheden af egenskabsestimater på det endelige design; liste og rangere egenskaberne, der er vigtigst fra procesdesignets synspunkt; og at etablere acceptable niveauer af nøjagtighed for egenskabsmodeller. Anvendelsen af denne metode er fremhævet gennem tre case studier, design af en ekstraktiv destillationproces, design af en short-path fordamper og design af et afsyringssystem til en vegetabilsk olie deodoriceringsproces.

Table of Contents

Preface ..... **i**

Abstract ..... **iii**

Resume på Danske ..... **v**

Table of Contents ..... **vi**

List of Tables ..... **x**

List of Figures ..... **xv**

Chapter 1. Introduction **1**

1.1 Modelling of Pure Component Properties for Chemical Product-Process Design and Analysis: Current Challenges and Issues 1

1.2 Objectives of Thesis 4

1.3 Thesis Organization 5

1.4 Dissemination of results obtained and presented in this thesis 6

Chapter 2. Theoretical Background **7**

2.1 Introduction 7

2.2 Group-Contribution^{+} (GC^{+}) Methods 9

2.3 Current State-of-the-Art 11

2.3.1 GC methods for the estimation of thermo-physical and transport-related properties of pure components 11

2.3.2 GC methods for the estimation of environmental-related properties of pure components 12

2.4 Current Challenges and Issues 13

2.4.1 Selection of quantity of property data used for parameter regression 13

2.4.2 Selection of suitable property prediction model function 14

2.4.3 Uncertainty analysis for evaluating reliability of property models 15

2.4.4 Themodynamic consistency and predictive power 16

2.4.5 Sensitivity of process design due to uncertainties of property estimates property estimates 17

2.5 Summary 18

Chapter 3. Methodology **19**
3.1 Basis for Selecting Marrero and Gani (MG) Method 19
3.2 Methodology for Performing Property Modeling and Analysis 20
3.2.1 Selection of quantity of property data for parameter regression 21
3.2.2 Selection of suitable property prediction model function 23
3.2.3 'Molecular structural similarity criteria' for improvement of property models 23
3.2.4 Parameter estimation and uncertainty analysis using maximum-likelihood estimation theory 24
3.2.5 Tests for thermodynamic consistency and predictive power 26
3.2.6 Sensitivity of chemical product and process design due to uncertainties of property estimates property estimates 27
3.3 Summary 29

Chapter 4. Results: Model Development and Verification **30**
4.1 Development of GC^+ Models for Thermo-Physical and Transport-Related Properties. 30
4.1.1 Database 30
4.1.2 Model performance statistics 41
4.1.3 Selection of quantity of data-set for parameter regression 41
4.1.4 Reliability of developed property models 41
4.1.5 Thermodynamic consistency and predictive power 42
4.1.6 Application example 47
4.1.7 Summary 50
4.2 Development of GC^+ Models for Thermo-Physical and Transport-Related Properties. 51
4.2.1 Database 55
4.2.2 Selection of suitable property model function 56
4.2.3 Model performance statistics 57
4.2.4 Selection of quantity of data-set for parameter regression 70
4.2.5 Application example 74
4.2.6 Summary 76
4.3 Further Improvement of GC Models Based on the 'Molecular Structural Similarity Criteria' Approach 77
4.3.1 Development of GC model for $\Delta_f H^o_{gas}$ with chemical Accuracy 77
4.3.2 Application of improved GC model for $\Delta_f H^o_{gas}$ 92

4.3.3 Application of 'molecular structure similarity criteria' to modeling of other pure component properties 95
4.3.4 Selection of minimum data-set for the parameter regression 96
4.3.5 Summary 97

Chapter 5. Results: Software Development 99
5.1 Introduction to ProPred 99
5.2 Inclusion of Developed GC$^+$ Based Property Models for Thermo-Physical and Transport-Related Properties in ProPred 100
5.3 Inclusion of Developed GC$^+$ Based Property Models for Environmental-Related Properties in ProPred 104
5.4 Inclusion of improved GC property models developed using molecular structure similarity criteria in the ProPred 106
5.5 Summary 108

Chapter 6. Results: Applications in Process Design 109
6.1 Case Study 1: Sensitivity Analysis of Design of Extractive Distillation Process due to Uncertainties of Property Estimates 110
6.2 Case Study 2: Sensitivity Analysis of Design of Short-Path Evaporator due to Uncertainties of Property Estimates 115
6.3 Case Study 3: Sensitivity Analysis of Design of De-acidification System due to Uncertainties of Property Estimates 122
6.4 Summary 129

Chapter 7. Conclusions and Future Work 130
7.1 Conclusions 130
7.2 Achievements and Contributions 132
7.3 Future Work and Recommendations 133

Nomenclature 134
References 136
Appendices 143
Appendix A. Group-contribution and atom-contribution tables for GC+ models for thermo-physical and transport-related properties 143

Appendix B. Group-contribution and atom-contribution tables for GC+ models for environmental-related properties 197

Appendix C. First-order, second-order, and third-order group- contributions for improved GC models for $\Delta_f H^o_{gas}$, $\Delta_{fus} H$, and T_c 253

Appendix D. List of molecular-group compounds 261

Appendix E. Mathematical model for short-path evaporator developed by Sales-Cruz and Gani (2006) 262

Appendix F. List of publications and the conference presentations related to this PhD project 264

List of Tables

Table 2.1	Representation of molecular structure of 5-Hydroxymethylindane using Marrero and Gani groups	9
Table 3.1	Occurrence matrix and experimental data of $\Delta_f H^o_{gas}$ for *n*-alkanes and *n*-alkenes	22
Table 4.1	Class-wise description of the CAPEC database used in the parameter regression	32
Table 4.2	Performance of MG method based models for thermo-physical and transport properties analysed using step-wise regression method	33
Table 4.3	Performance of CI method based property models for thermo-physical and transport-related properties	37
Table 4.4	Values of the universal constants (additional adjustable parameters)	40
Table 4.5	Effect of quantity of experimental data on the quality of parameter estimation	41
Table 4.6	Comparison of model prediction error with reported average measurement error	42
Table 4.7	Estimation of T_b of Phenanthrene, 2-methyl-	47
Table 4.8	Covariance matrix $COV(P^*)$ with dimensions (7×7)	48
Table 4.9	Local sensitivity $J(P^*)$ with dimensions (1×7) of T_b model with respect to model parameters	48
Table 4.10	Definition of environmental-related properties considered in this work	52
Table 4.11(a)	Description of the US EPA data-sets used for the regression purpose	55
Table 4.11(b)	Description of the USEtox^{TM} data-sets used for the regression purpose	55
Table 4.12	Performance of MG method based property models for environmental-related properties analysed using step-wise regression method	59
Table 4.13	Performance of CI method based property models for environmental-related properties	64

Table 4.14	Values of Universal Constants (Additional Adjustable Parameters)	69
Table 4.15(a)	Performance of model for Oral Rat LD_{50} based on different combinations of training sets and test sets	72
Table 4.15(b)	Performance of model for fathead minnow 96-h LC_{50} based on different combinations of training sets and test sets	72
Table 4.15(c)	Model performance for emission to Urban Air (Carcinogenic) based on different combinations of training and test sets	73
Table 4.15(d)	Model performance for emission to Urban Air (Non-Carcinogenic) based on different combinations of training and test sets	73
Table 4.16	Estimation of $LogW_s$ of Benzo[a]pyrene	74
Table 4.17	Covariance matrix $COV(P^*)$ with dimensions (7×7)	75
Table 4.18	Local sensitivity $J(P^*)$ with dimensions (1×7) of $LogW_s$ model with respect to the model parameters	75
Table 4.19	Class-wise description of the data-set of $\Delta_f H^o_{gas}$	78
Table 4.20	Classification of data-set of $\Delta_f H^o_{gas}$ based on observed deviation	81
Table 4.21	List of new third-order groups defined for the Marrero and Gani GC-method with sample assignments	82
Table 4.22	Model performance statistics for the developed GC model for $\Delta_f H^o_{gas}$	84
Table 4.23	Performance of improved GC model for $\Delta_f H^o_{gas}$ for various classes of components	85
Table 4.24	Distribution of data-points based on the range of absolute deviations from the developed GC model for $\Delta_f H^o_{gas}$	86
Table 4.25	Comparison of prediction performance between the improved GC model and the Benson's method (1993)	87
Table 4.26	Comparison of the distribution of data-points based on the range of absolute deviations from the developed GC model and from the Benson's method (1993)	88
Table 4.27	Comparison of prediction performance between the developed GC model and the Benson's method using the ASTM CHETAH® software and the NIST® software for selected organic components	90

Table 4.28	Prediction performance for 20 organic compounds not included in the parameter regression step	92
Table 4.29	Prediction of $\Delta_f H^o_{gas}$ for 2-Methylquinoline	93
Table 4.30	Covariance matrix $COV(P^*)$ for the groups listed in Table 4.28	94
Table 4.31	Local sensitivity ($J(P^*)$ of $\Delta_f H^o_{gas}$ model	94
Table 4.32	Comparison of performance of GC models for $\Delta_{fus}H$ and T_c	95
Table 4.33	Chi-squared reduced statistics for the developed and previous version of GC models for $\Delta_f H^o_{gas}$, $\Delta_{fus}H$ and T_c	96
Table 4.34	Performance statistics for the total data-set, for the minimum data-set and for additional data-points used for the model validation.	97
Table 6.1	Base case data for extractive distillation column (Gil et al., 2 009)	111
Table 6.2	Base case data for entrainer recovery column (Gil et al., 2009)	111
Table 6.3	Problem data for the short-path evaporator (Cruz and Gani, 2006)	116
Table 6.4(a)	Composition of acylglycerides in CPO (Gee, 2007)	123
Table 6.4(b)	Composition of FFA's and other components in CPO (Gee, 2007)	123
Table 6.5	Stream summary for de-acidification process from PRO/II®	125
Table A1	Performance of MG method based property models for thermo-physical and transport-related properties analysed using simultaneous regression method	144
Table A2	MG method based property models analysed using step-wise regression method: First-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, ω, and V_{mo}	148
Table A3	MG method based property models analysed using step-wise regression method: Second-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, ω ,and V_{mo}	158
Table A4	MG method based property models analysed using step-wise regression method: Third-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, ω ,and V_{mo}	166

Table A5 MG method based property models analysed using simultaneous regression method: First-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, AiT, ω, and V_{mo} 172

Table A6 MG method based property models analysed using simultaneous regression method: Second-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, AiT, ω, and V_{mo} 182

Table A7 MG method based property models analysed using simultaneous regression method: Third-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, AiT, ω, and V_{mo} 190

Table A8 CI method based property models: atom contributions and model constants for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, ω, and V_{mo} 196

Table B1 Performance of MG Method Based Property Models Analysed by Simultaneous Regression Method 198

Table B2 MG Method Based Property Models Analysed Using Step-Wise Regression Method: First-Order Groups and their Contributions for the Environmental-Related Properties 203

Table B3 MG Method Based Property Models Analysed Using Step-Wise Regression Method: Second-Order Groups and their Contributions for the Environmental-Related Properties 213

Table B4 MG Method Based Property Models Analysed Using Step-Wise Regression Method: Third-Order Groups and their Contributions for the Environmental-Related Properties 221

Table B5 MG Method Based Property Models Analysed Using Simultaneous Regression Method: First-Order Groups and their Contributions for the Environmental-Related Properties 227

Table B6 MG Method Based Property Models Analysed Using Simultaneous Regression Method: Second-Order Groups and their Contributions for the Environmental-Related Properties 237

Table B7	MG Method Based Property Models Analysed Using Simultaneous Regression Method: Third-Order Groups and their Contributions for the Environmental-Related Properties	245
Table B8	CI Method Based Property Models: Atom Contributions and Model Constants for the Environmental-Related Properties	251
Table C1	The total list of first-order groups and their contributions C_i for the improved GC models for T_c, $\Delta_f H^o_{gas}$, and $\Delta_{fus}H$	254
Table C2	The total list of second-order groups and their contributions D_j for the improved GC models for T_c, $\Delta_f H^o_{gas}$, and $\Delta_{fus}H$	257
Table C3	The total list of third-order groups and their contributions E_k for the improved GC models for T_c, $\Delta_f H^o_{gas}$, and $\Delta_{fus}H$	259
Table D1	List of molecular-group compounds for which developed GC model for $\Delta_f H^o_{gas}$ is not applicable	261

List of Figures

Figure 1.1	Important issues related to GC based property modeling	3
Figure 2.1	Illustration of concept of the GC method	8
Figure 2.2	Molecular structure of Diazene, diphenyl- (CAS No. 103-33-3)	9
Figure 2.3	Plot of experimental values of T_m versus carbon number of pure components	15
Figure 2.4	Plot of T_c/T_b for *n*-alkanes versus their increasing carbon number	17
Figure 3.1	Application of molecular structural similarity criteria	23
Figure 3.2	Methodology for performing the sensitivity analysis for process design	27
Figure 4.1	T_b versus T_c for *n*-alkanes	43
Figure 4.2	Plots for: (a) T_c/T_b ratio for a wide range of components in the data set; (b) Experimental values of T_c/P_c versus predicted values of T_c/P_c.	44
Figure 4.2	Plots for: (c) Predicted P_c of *n*-alkanes; (d) Predicted Z_c of *n*-alkanes.	46
Figure 4.3	Experimental values of T_b and the calculated 95% confidence intervals for the pure components in the data-set	49
Figure 4.4	Plot of molecule types versus their experimental values of –Log LC_{50}(FM)	56
Figure 4.5	Experimental values of $LogW_s$ and the calculated 95% confidence intervals versus data-set of $LogW_s$	76
Figure 4.6(a)	Plot of $\Delta_f H^o_{gas}$ versus carbon number of pure components	79
Figure 4.6(b)	Plot of $\Delta_f H^o_{gas}$ versus carbon number of pure components	79
Figure 4.7	Plot of experimental and predicted values of enthalpy of formation versus carbon number of pure components	91
Figure 5.1	Starting window in ProPred	100
Figure 5.2	ProPred menu bar	100
Figure 5.3	A screenshot of 'Marrero and Gani' page of ProPred	101

Figure 5.4	A screenshot of group assignments section in 'Marrero and Gani' page	102
Figure 5.5	A screenshot of molecular structure involving a missing group (PH_2)	103
Figure 5.6	A screenshot of predicted group-contributions values for the missing groups	103
Figure 5.7	A screenshot of predicted properties using the GC^+ approach	104
Figure 5.8	A screenshot of estimation of environmental-related properties needed for the WAR algorithm	105
Figure 5.9	A screenshot of estimation of environmental-related properties needed for the USEtox model	105
Figure 5.10	A screenshot of estimation of enthalpy of formation using improved GC model	107
Figure 5.11	A screenshot of new third-order group displayed in ProPred	107
Figure 6.1	Steady state process model for extractive distillation process in PRO/II	112
Figure 6.2	(a) Sensitivity of R due to errors of P_{vap}; (b) Sensitivity of N due to errors of P_{vap}; (c) Sensitivity of Ns due to errors of P_{vap}; (d) Sensitivity of Q_r due to errors of $\Delta_{vap}H$; (e) and (f) Sensitivity of D due to errors of ρ and σ resp.; (g) Sensitivity of R due to errors of T_c	114
Figure 6.3(a)	Sensitivity of product flow rate due to uncertainties of predicted P_{vap}	117
Figure 6.3(b)	Sensitivity of product flow rate due to uncertainties of T_c	118
Figure 6.3(c)	Sensitivity of product flow rate due to uncertainties of P_c	118
Figure 6.3(d)	Sensitivity of product flow rate due to uncertainties of ω	119
Figure 6.3(e)	Sensitivity of product flow rate due to uncertainties of predicted C_p	120
Figure 6.3(f)	Sensitivity of product flow rate due to uncertainties of predicted k	120
Figure 6.3(g)	Sensitivity of product flow rate due to uncertainties of predicted ρ	120
Figure 6.3(h)	Sensitivity of product flow rate due to uncertainties of predicted μ	121
Figure 6.4	Steady state model of de-acidification process in PRO/II®	124

Figure 6.5(a)	Sensitivities of PFAD flow from cold scrubber to uncertainties of predicted P_{vap}	126
Figure 6.5(b)	Sensitivities of FFA's to vacuum system due to uncertainties of predicted P_{vap}	127
Figure 6.5(c)	Sensitivities of composition of Toco's to uncertainties of predicted P_{vap}	127
Figure 6.5(e)	Sensitivity of diameter of the cold scrubber to uncertainties of predicted ρ	128
Figure 6.5(f)	Sensitivity of diameter of the hot scrubber to uncertainties of predicted ρ	128
Figure 6.5(g)	Sensitivities of PFAD flow from cold scrubber to uncertainties of T_c	128
Figure E.1	Balance volumes for the short-path evaporator. The balance volume 1 is for the mass and energy balance; the balance volume 2 is for the momentum balance (Adapted from Cruz and Gani, 2006)	262

Figure 6.5(a)	Sensitivities of PFAD flow from cold scrubber to uncertainties of predicted P_{vap}	126
Figure 6.5(b)	Sensitivities of FFA's to vacuum system due to uncertainties of predicted P_{vap}	127
Figure 6.5(c)	Sensitivities of composition of Toco's to uncertainties of predicted P_{vap}	127
Figure 6.5(e)	Sensitivity of diameter of the cold scrubber to uncertainties of predicted ρ	128
Figure 6.5(f)	Sensitivity of diameter of the hot scrubber to uncertainties of predicted ρ	128
Figure 6.5(g)	Sensitivities of PFAD flow from cold scrubber to uncertainties of T_c	128
Figure E.1	Balance volumes for the short-path evaporator. The balance volume 1 is for the mass and energy balance; the balance volume 2 is for the momentum balance (Adapted from Cruz and Gani, 2006)	262

Chapter 1. Introduction

1.1 Modeling of Pure Component Properties for Chemical Product-Process Design and Analysis: Current Challenges and Issues

In chemical process design, one determines the conditions (temperature, pressure, and/or composition) at which the properties (such as solubility, vapor pressure etc.) of an identified set of pure components match the design objectives. In chemical product design, one determines the pure components that exhibit certain desirable or specified properties/behavior (Constantinou and Gani, 1996). For example, pure component properties as well as mixture properties (calculated using pure component properties and suitable mixing rules) are needed to study the behavior of the chemical product (such as the solubility of a drug molecule in water), dimensioning of equipments (by using the densities of chemicals present in a tank), study the physical state of a chemical (melting point and/or normal boiling point of a chemical to identify its phase) and many more. Therefore, the knowledge of physical and thermodynamic properties of pure components is a basic requirement for performing tasks related to process design, simulation, and optimization as well as design of chemicals based products. The environmental-related properties such as, fathead minnow 96-h LC_{50}, oral rat LD_{50}, etc., on the other hand, play an important role in performing sustainability analysis using a suitable technique such as the Waste Reduction (WAR) algorithm (Young and Cabezas, 1999) and the USEtox^{TM} model. For example, the WAR algorithm is used together with a chemical process simulator to quantify the impact that the processes would have on the environment and thereby to identify sustainable processing paths and design alternatives. The USEtox^{TM} model (Rosenbaum et al., 2011) is an environmental model for characterization of human and ecotoxicological impacts in life cycle impact assessment (LCIA) and Comparative Risk Assessment (CRA) and is designed to describe the fate, exposure and effects of chemicals. The USEtox^{TM} model calculates characterization factors for carcinogenic and noncarcinogenic impacts based on the chemical's properties related to emissions to urban air, rural air, fresh water, sea water, agricultural soil, and natural soil. Therefore, the basis for performing sustainability analysis of chemical processes using the WAR and the USEtox^{TM} model is a set of environmental-related properties of chemicals involved in that process.

An important limiting factor in performing chemical product-process design and analysis, however, is the availability of reliable values of properties of pure components. There are three ways in which a property user can obtain the information of needed properties: (i) by retrieving the property information available in databases/open literature; (ii) by performing

laboratory measurements for the needed properties; and (iii) by employing suitable property prediction methods. A key limitation associated with the use of databases is the limited number of pure components (and sometimes limited number of properties) stored in the database. While use of experimentally measured property values is desirable, laboratory measurements may be time consuming, expensive, and sometimes may not even be feasible. In addition, the gap between the number of pure components registered in *Chemical Abstracts* and the available experimental values of properties of pure components is continuously increasing. In such a situation, the most convenient and practical approach has been to employ property prediction models to estimate necessary pure component properties from their molecular structures, at least in the early stages of product-process design.

Property prediction implies the use of mathematical models for the estimation of the needed properties. Depending on the type of the property, the required prediction accuracy, type of application etc., these mathematical models may be simple polynomial functions representing various pure component properties, a non-linear set of algebraic equations representing cubic equations of state, or a very large set of differential-algebraic system representing behavior at the atomic scale (Kontogeorgis and Gani, 2004). An ideal property prediction model developed for the estimation of pure component properties would (1) provide accurate and reliable estimates of needed properties for pure components, (2) require a minimum of input data for the property estimation, (3) provide an indication of prediction errors (that is, uncertainties of predicted property values), (4) provide large application range and good predictive capability, and (5) require less computational needs and time. For the prediction of pure component properties, methods such as, group-contribution (GC), quantitative structure-property relationship (QSPR), and *ab initio* quantum mechanics based methods are widely employed. *Ab initio* methods, including the popular Density Function Theory (DFT) based methods, involve solution of the quantum mechanical Schrödinger's equation in a rigorous way. These methods yield good prediction accuracy for very small molecules (Peterson et al., 2012; Chan et al., 2011; Karton et al., 2011). As these methods have been mainly developed and tested on really small molecules, their proven applicability is very limited. Moreover, calculations for *ab initio* methods are computationally highly demanding and as such are not readily usable by the chemist and technologist in daily practice. In QSPR based methods, the basic idea is to find a relationship between the structure of a pure component expressed in terms of constitutional, topological, geometrical, electrostatic and quantum-chemical descriptors and the physical property of interest. QSPR methods yield explicit or implicit correlations that can be used to estimate properties of those pure components that are structurally similar to the components present in the training set (Bunz et al., 1999). The correlative power of these approaches has been demonstrated in many cases. However, the extrapolative ability of these methods is limited, especially when these methods are used outside the training set. In GC methods, the property of a pure component is a function of

structurally dependent parameters, which are determined as a function of the frequency of groups representing the pure component and their contributions. In principle, it is an additive method, where the contributions of each group towards a property are summed to obtain the property value. Among these three above mentioned property prediction methods, the GC based property prediction methods are commonly used since these are simple, computationally undemanding, and can be used 'on the fly' by the process engineer or chemist. Therefore, in this work, the development and analysis of property prediction models for pure component properties based on the GC approach have been considered.

Though, GC methods have been frequently used by practitioners for performing chemical product-process design and analysis, a number of issues remained challenging for the development of property models. Figure 1.1 shows the most common and important issues related to these namely: (i) selection of the quantity of property data used in the parameter regression step; (ii) selection of the most appropriate form of the property model function; and (iii) accuracy, thermodynamic consistency, predictive power, and application range of GC models.

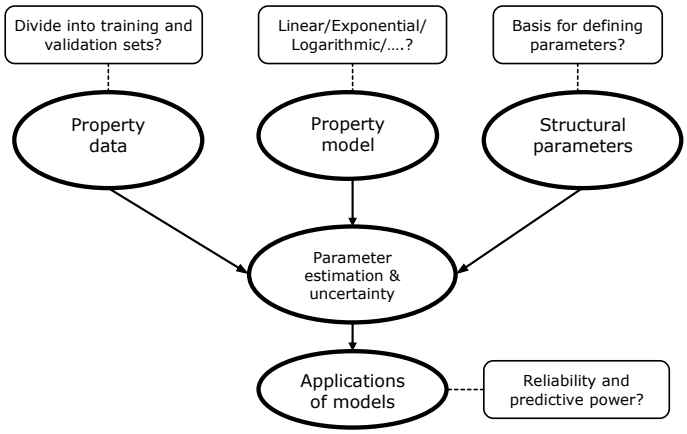


Figure 1.1 Important issues related to GC based property modeling

In the development of a GC based property model, it is often preferred to use some portion of the data-set for training purpose and use remaining portion for validating the reliability of the predictions from the GC model. The division of the data-set into training set and validation set may not allow regression of some GC model parameters (that is, group-contributions) due to the lack of necessary data that is made unavailable while making the validation set using a random selection technique. This affects the application range of the GC model. Further, the

prediction accuracy of GC model will be lower since lower quantity of the property data (and hence, less information) is used in the parameter regression step.

The selection of the property model function (linear/exponential/logarithmic etc.) decides the accuracy, consistency, and predictive power of GC based property models. Therefore, the property model developer needs to consider the information of pure component property and analysis of the property data while selecting the most appropriate form of the property model function.

Other key issues associated with the use of GC models are accuracy, application range, thermodynamic consistency, and their predictive power. A common approach for evaluating the performance of GC models has been an analysis of statistical performance indicators such as coefficient of determination, standard deviation, average absolute error, and average relative error obtained from the parameter regression step. These indicators provide measures of quality (reliability) of property prediction models. However, for assessing the quality and reliability of predicted property values themselves, it is also necessary to know the uncertainties (for example, 95% confidence intervals) of the predicted property values individually. The quantification of uncertainties of property estimates allows the property user to perform better-informed chemical product-product design by taking into account the effects of uncertainties of the property estimates. To this end, sensitivity analysis of product-process design due to uncertainties of property estimates can provide useful information on the effects of these uncertainties on the quality of the design.

Finally, it is important to perform tests to verify the thermodynamic consistency and predictive power of GC models, especially when property models are to be used outside their proven range. This is particularly true in the design and analysis of bio-chemical, food, and pharmaceutical processes, which involve molecules that are large, complex, or entirely new (that is, molecules yet to be synthesized).

1.2 Objectives of Thesis

Motivated by the need and importance of accurate and reliable predictions of properties of pure components and uncertainties of predicted property values in product-process design and analysis, this work aims to achieve the following objectives:

a) Develop a systematic methodology for performing property modeling of pure component properties. The methodology is to include: (i) selection of quantity of data-set for parameter regression; (ii) selection of suitable property prediction model

function; (iii) uncertainty analysis of property models; and (iv) tests for thermodynamic consistency and predictive power of property models.

b) Develop GC^+ based property models using the systematic methodology developed in **(a)** to provide accurate and reliable predictions of pure component properties together with uncertainties of predicted property values. The property modeling and uncertainty analysis of thermo-physical properties, transport-related properties, and environmental-related properties of pure components have been considered in this work.

c) Develop and apply the 'molecular structural similarity criteria' based approach to make efficient use of knowledge of properties in the development of property models to achieve the targeted level of prediction accuracy. Through 'molecular structural similarity criteria' based approach, the development of a GC model for enthalpy of formation of pure components is aimed to achieve the targeted prediction accuracy of ± 2 kJ/mol (*chemical accuracy*).

d) Implement the developed property models into the ProPred, a property estimation toolbox of Integrated Computer Aided System (ICAS®) developed at CAPEC, DTU so that it is fast and convenient for the user to obtain pure component properties.

e) Develop and illustrate a systematic framework for performing analysis of sensitivity of process design due to uncertainties of property estimates to evaluate the effects of uncertainties of property estimates on the final design using three case studies.

1.3 Thesis Organization

This Ph.D thesis is organized into seven chapters including the current chapter (Introduction). In the current chapter, an overview of challenges and issues related to property modeling and specific needs to overcome these challenges are briefly reviewed. Based on the needs identification, the objectives of the PhD thesis are discussed. Chapter 2 presents overview of the theoretical background of different GC based property prediction models currently used and a deeper analysis of key issues involved in the property modeling. The effects of uncertainties of property estimates on the design of several important chemical processes are also briefly reviewed in Chapter 2. Chapter 3 presents description of the developed methodology for performing property modeling and uncertainty analysis and a framework for performing sensitivity analysis of process design due to uncertainties of property estimates. The development and analysis of GC^+ based property models for the estimation of thermo-physical, transport-related, and environmental-related properties is discussed in Chapter 4. The application of "molecular structural similarity criteria" based approach to improve the performance of GC models for enthalpy of formation, enthalpy of fusion, and critical temperature is also discussed. Chapter 5 mainly focuses on the implementation of developed property models into the ProPred. In Chapter 6, three case studies highlighting the application

of developed methodology for performing sensitivity analysis of process design due to uncertainties of property estimates are presented. Finally, the thesis is concluded with chapter 7 summarizing contributions from this work and a brief discussion on directions for future developments of this work.

1.4 Dissemination of results obtained and presented in this thesis

The developed methods for performing property modeling and uncertainty analysis and their application to the development of GC^+ based property models for prediction of pure component properties have been discussed at 13 international conferences through presentations. Furthermore, the developed methods, results of property modeling and uncertainty, and application of developed property models in process design have been disseminated through publications in scientific journals and conference proceedings. In particular, the methodology for performing property modeling and uncertainty analysis discussed in chapter 3 and its application to development of GC^+ models for the estimation of thermo-physical properties discussed in chapter 4 has been presented in Hukkerikar et al. (2012a). The development and application of the method for performing sensitivity analysis of process design due to uncertainties of property estimates discussed in chapter 6 has been presented in Hukekrikar et al. (2012b). The property modeling and uncertainty analysis of GC^+ models for estimation of environmental-related properties discussed in chapter 4 has been presented in Hukkerikar et al. (2012c). The "molecular structural similarity criteria" based approach and its application to development of GC models for the estimation of pure component properties with desired prediction accuracy as discussed in chapter 4 has been presented in Hukkerikar et al. (2013a) and in Hukkerikar et al. (2013b). Furthermore, a review on the available GC^+ models for the estimation of properties of lipid components and recommendations for further improvements has been presented in Cunico et al. (2013c).

Chapter 2. Theoretical Background

2.1. Introduction

Property prediction models for pure components can be classified into models for predicting single value properties of pure components (such as normal boiling point, critical constants, normal melting point etc.), and models for predicting temperature dependent properties of pure components (such as liquid vapor pressure, liquid heat capacity, liquid viscosity etc.). The estimated single value properties can be classified into primary properties and secondary properties. The primary properties are those properties that are estimated using only the molecular structural information (for example, functional groups) of pure components. The secondary properties, on the other hand, are the properties that are estimated using the molecular structural information and other variables/primary properties of pure components (Constantinou and Gani, 1996). Further, the single value pure component properties are classified into: (i) thermo-physical properties such as, normal boiling point, critical constants, enthalpy of formation, etc.; (ii) transport-related properties such as, viscosity, thermal conductivity, surface tension, etc.; and (iii) environmental-related properties such as fathead minnow 96-h LC_{50}, oral rat LD_{50}, aqueous solubility, emission to urban air (carcinogenic), emission to urban air (non-carcinogenic), etc. In the following section, since the focus of this work has been on the development and analysis of GC based property models for estimation of single value pure component properties, an overview of the available GC methods and further discussion on the key issues involved in the property modeling are discussed.

In the GC method, the molecular structure of a pure component is decomposed into building blocks and the property of that molecule is estimated by the summation of the contributions of these building blocks. The building blocks are referred as functional groups. These methods are fast, simple to use, and are predictive in nature because the same functional group can be used to represent the molecular structure of more than one pure component. For example, to calculate the standard heat of formation of a pure component, only a simple linear summation of the contributions of the structural groups representing the molecular structure of that pure component is needed. An example illustrating the concept of the GC method is shown in Figure 2.1. The molecular structure of 1-Hexanol is represented by 1 CH_3, 5 CH_2, and 1 OH groups while molecular structure of Tetradecanoic acid is represented by 1 CH_3, 12 CH_2, and 1 COOH groups. In the GC approach, the contributions of the CH_3 and CH_2 group to the pure component properties of 1-Hexanol and Tetradecanoic acid are considered to be the same.

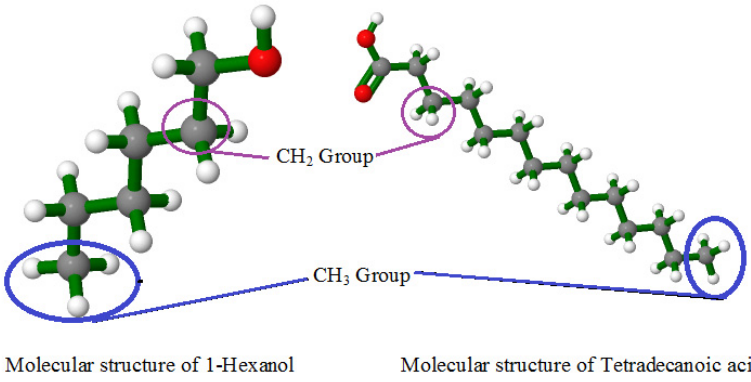


Figure 2.1 Illustration of concept of the GC method

The simplest form of a GC method employs first-order groups for the estimation of properties of pure components as shown in Eq. (2.1).

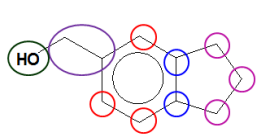
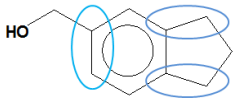
$$f(X) = \sum_i N_i C_i \tag{2.1}$$

Where, the function $f(X)$ is a function of property X. In Eq. (2.1), C_i is the contribution of the first-order group of type-i that occurs N_i times. For the estimation of properties of pure components, most of the currently-used methods employ only first-order groups. These methods work well for simple molecules but can yield significantly large prediction errors if used for the property estimation of large and complex structured molecules. To account for interactions between groups, e.g., more complex interactions in heterocyclic species, second-order (Constantinou and Gani, 1994) and third-order groups (Marrero and Gani, 2001) have been introduced. Further description on the details of first-order, second-order and third-order Marrero and Gani groups has been given in section 3.1 of chapter 3. The property prediction model employing the Marrero and Gani (2001) GC method (also known as the MG method) has the form,

$$f(X) = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k E_k O_k \tag{2.2}$$

The definition of $f(X)$ is specific for each property X. For example, the $f(X)$ for the property model of normal melting point (T_m) of pure components is $\exp(T_m/T_{mo})$ where, T_{mo} is an additional adjustable parameter of the GC model. In Eq. (2.2), D_j is the contribution of the second-order group of type-j that occurs M_j times. E_k is the contribution of the third-order group of type-k that has O_k occurrences in the molecular structure of a pure component. An example showing representation of the molecular structure of the component: 5-Hydroxymethylindane (CAS No. 51632-06-5) using first-order, second-order and third-order groups of the MG method is shown in Table 2.1.

Table 2.1 Representation of molecular structure of 5-Hydroxymethylindane using Marrero and Gani groups

First-order groups[a]/ their occurrences	Second-order groups[a]/ their occurrences	Third-order groups[a]/ their occurrences
aCH /3 aC /2 aC-CH_2 /1 OH /1 CH_2 (cyc) /3	aC-CH_n-OH (n in 1..2) /1	aC-$CH_{ncyclic}$ (fused rings) (n in 0..1) /2 AROM.FUSED[2]S^2 /1
		

[a] The colored circles in the drawing of the structure of 5-Hydroxymethylindaneshow the listed groups in respective colors. The S^2 in the group AROM.FUSED[2]S^2 represents the position of branched fragment in the fused ring structure.

2.2 Group-Contribution$^+$ (GC$^+$) Methods

The application range of the developed GC model largely depends on the data-set of pure components considered in the parameter regression step, structural parameters (that is, groups) used to represent the molecular structure of pure components, and the selected property model function, $f(X)$. Let us consider the molecular structure of Diazene, diphenyl- (CAS No. 103-33-3) (see Figure 2.2).

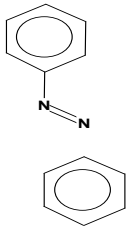


Figure 2.2 Molecular structure of Diazene, diphenyl- (CAS No. 103-33-3)

Let us assume that it is required to estimate critical temperature of Diazene, diphenyl-. If the contribution of any of the involved groups in the molecular structure of Diazene, diphenyl-, say N=N group, is not available in the parameter value table of the GC model employed to

estimate the critical temperature, then the estimation of critical temperature of Diazene, diphenyl is not possible. Such groups-contributions are referred to as missing group-contributions. One way to address this limitation with the group contribution methods is to use new experimental data in the parameter regression so that the missing group-contributions can be estimated. This procedure will require time and resources and hence this option is not convenient for the model user (Gani et al., 2005). Further, for the case where the molecular structure of the pure component whose properties need to be estimated is not completely described by any of the available groups (that is, missing groups) of a particular GC method, the property estimation using that GC method is not possible. To overcome these limitations, Gani et al. (2005) combined the atom connectivity index (CI) method with the GC method to create missing groups and/or to predict their contributions. This combined approach has led to the development of group contribution^{+} (GC^{+}) method of a wider application range than before since the missing groups and their contributions can now be easily predicted without the need of additional experimentally measured data needed in for the parameter regression. Motivated by the need and usefulness of GC^{+} based approach, the GC^{+} approach has been extended to modeling of polymer properties (Satyanarayana et al., 2007; Satyanarayana et al., 2009). Furthermore, the GC^{+} approach has been employed in the property modeling of Hansen solubility parameters (Modaressi et al., 2008), as well as in modeling of liquid viscosity and liquid surface tension of pure components (Conte et al., 2008). The CI method employs the following model for the estimation of pure component properties.

$$f(X) = \sum_i a_i A_i + b\left({}^{v}\chi^{0}\right) + 2c\left({}^{v}\chi^{1}\right) + d \tag{2.3}$$

Where a_i is the contribution of the atom of type-i that occurs A_i times in the molecular structure, ${}^{v}\chi^{0}$ is the zeroth-order (atom) valence connectivity index, ${}^{v}\chi^{1}$ is the first-order (bond) valence connectivity index, b and c are adjustable parameters, and d is a universal parameter. Please note that $f(X)$ of models in the MG method and in the CI method, (i.e. left hand side of Eq. (2.2) and Eq. (2.3)) has the same functional form for a particular pure component property X and the values of universal constants for the CI models are the same as those for the MG method. High accuracy in the prediction of X cannot be expected from this model since, with only a few parameters, a large set of components is going to be represented. Greater accuracy can be obtained by adding higher-order connectivity indices (He and Zhong, 2003; Lin et al., 2005). However, for the purpose of creating missing groups and missing group contributions, only the first two connectivity indices are sufficient (Gani et al., 2005). First, the parameters, ${}^{v}\chi^{0}$ and ${}^{v}\chi^{1}$ for the missing groups as well as for the entire molecule are calculated using the rules described by Gani et al. (2005). Once these indices are calculated, following CI model equations are applied to the missing groups to compute $f(X_m)$ and $f(X^{*})$.

$$f\left(X_m\right)=\sum_i a_{m,i}A_{m,i}+b\left({}^{v}\chi^{0}\right)_m+2c\left({}^{v}\chi^{1}\right)_m \tag{2.4}$$

$$f\left(X^*\right)=\left(\sum_m n_m f\left(X_m\right)\right)+d \tag{2.5}$$

Where m is the number of different missing groups and n_m indicates the number of times a missing group appears in the molecule. Finally, the value of property X is estimated using the following GC$^+$ model equation:

$$f\left(X\right)=\sum_i N_i C_i+f\left(X^*\right)+\sum_j M_j D_j+\sum_k E_k O_k \tag{2.6}$$

2.3 Current State-of-the-Art

2.3.1 GC methods for estimation of thermo-physical and transport-related properties of pure components

For the estimation of thermo-physical and transport-related pure component properties, various GC methods have been developed by many authors. Lydersen (1955), Ambrose (1978), and Klincewicz and Reid (1984) developed GC methods for the prediction of critical constants. The first two GC methods use group-contributions obtained by analyzing incremental changes in physical properties within homologous series while the third method uses a least square regression using the property data from Ambrose (1978). Joback and Reid (1987) developed GC based models for the prediction of 9 pure component properties including critical constants, normal boiling point and melting point among others. There are some limitations of this method which include: (i) small sized data-sets used for regressing the model parameters; (ii) the method employs only 41 functional groups, which oversimplifies the molecular structure thus making several types of isomers indistinguishable; and (iii) issue related to the selection of appropriate form of model functions for some properties. Marrero and Pardillo (1999) developed a group-interaction (GI) approach based method to predict normal boiling point and critical constants of pure components. This method considers the contributions of interactions between the groups instead of contributions of groups. To overcome some of the limitations of previous methods, more sophisticated GC methods based on the multi-level property estimation approach have been developed by Constantinou and Gani (1994), Marrero and Gani (2001), and Dalmazzone et al. (2006). In Constantinou and Gani GC method, the property estimation is performed at two levels. The basic level has contributions from first-order functional groups and the next level has second-order groups, which are defined based on the principle of conjugation. The properties that can be predicted using the Constantinou and Gani GC method are normal

boiling point, normal melting point, critical pressure, critical temperature, critical volume, standard enthalpy of vaporization at 298 K, standard Gibbs energy, standard enthalpy of formation at 298 K, acentric factor and liquid molar volume. Marrero and Gani (2001) have developed a GC method in which the property estimation is performed in three levels. The first level has a large set of simple groups that is able to partially capture proximity effects, but is unable to distinguish between isomers. For this reason, the first level of estimation is intended to deal with simple and mono-functional components. The second level permits a better description of polyfunctional components and differentiation among some isomers. Second-order groups are, however, unable to provide a good representation of components containing more than one ring as well as, in some cases, open-chain poly-functional components with more than four carbon atoms in the main chain. Thus, a further level is required to provide a better description for these types of components. This is accomplished by the introduction of third-order groups, which intend to represent the molecule at the third level of approximation. The third level allows estimation of complex heterocyclic and large (C=7 to 60) poly-functional acyclic components. The MG method has been employed to model a wide range of properties of pure components. Kolska et al. (2005) developed a MG method based property models for the estimation of enthalpy of vaporization (298 K) and enthalpy of vaporization (at boiling point); Modarresi et al. (2008) developed MG method based property models for the estimation of Hansen solubility parameters; and Conte et al. (2009) developed a MG method based property models for the estimation of liquid viscosity and surface tension of pure components. Nonnoolal et al. (2004; 2007; 2009) have developed GC methods to estimate the normal boiling point, critical constants, and the liquid viscosity of pure components. These authors considered group-interaction terms, additional correction terms, and new groups within their previous version of GC model to achieve improvement in the property estimation for those pure components that involve strongly associating groups (that is, those groups which do not follow the group additivity principle). Recently, Tabernero et al. (2012) developed a GC method for the estimation of enthalpy of fusion of polycyclic aromatic/aliphatic hydrocarbons (PAH) containing pure components. These authors, however, employed only 41 first-order functional groups to represent the molecular structures thus limiting the application range of the GC method to certain types of PAH containing pure components.

2.3.2 GC methods for estimation of environmental-related properties of pure components

A review article by Boethling et al. (2004) discusses available experimental data sources and various estimation methods including GC methods, QSPR methods, and correlation equations for obtaining values of environmental-related properties of chemicals. For the estimation of fathead minnow 96-h LC_{50} and aqueous solubility, various GC methods have been developed. Martin and Young (2001) developed a GC method to correlate the acute toxicity (96-h LC_{50})

to the fathead minnow using 397 organic chemicals. Casalegno et al. (2005) used a diatomic fragment approach based GC method to correlate the acute toxicity (96-h LC_{50}) of 607 organic chemicals. For the estimation of aqueous solubility, Marrero and Gani (2002) developed a GC method using a three-level parameter estimation approach with a dataset of 2087 organic chemicals used for the parameter regression purpose. There are several other GC methods available for the estimation of aqueous solubility (Klopman and Zhu, 2001; Kühne et al., 1995). For the estimation of oral rat LD_{50} and bio-concentration factor (BCF), the more common approach has been to employ correlation equations. For example, bio-concentration factor for a chemical is estimated using known value of its octanol/water partition coefficient. Martin et al. (2008) have developed a hierarchical clustering technique to predict a variety of endpoints, including oral rat LD_{50}, BCF, aqueous solubility, fathead minnow 96-h LC_{50} that combines group contributions with descriptors from graph theory. To the best of our knowledge, there are no GC methods reported in the literature for the estimation of following environmental-related properties: permissible exposure limit (OSHA-TWA), global warming potential, photochemical oxidation potential, ozone depletion potential, acidification potential, emission to urban air (carcinogenic and non-carcinogenic), emission to continental rural air (carcinogenic and non-carcinogenic), emission to continental fresh water (carcinogenic and non-carcinogenic), emission to continental sea water (carcinogenic and non-carcinogenic), emission to continental natural soil (carcinogenic and non-carcinogenic), and emission to continental agricultural soil (carcinogenic and non-carcinogenic).

2.4 Current Challenges and Issues

The key challenges and issues related to pure component property modeling are discussed in detail in following sub-sections 2.4.1 to 2.4.5. In chapter 3, a systematic methodology for property modeling and analysis of pure component properties is discussed to address and solve these challenges and issues.

2.4.1 Selection of quantity of property data used for the parameter regression

The property data play an important role in the development of property prediction models and hence the data-sets of experimentally measured property values need to be carefully analyzed and selected before performing the parameter regression. While developing a GC based model, the measured data-set of property values is often divided into the training set and the validation set. For example, Nannoolal et al. (2004) used a validation set consisting of 199 data-points (out of total 2850 data-points) for testing the prediction accuracy of their GC model for normal boiling point. Klopman and Zhu (2001) used a validation set consisting of 120 data-points (out of 1168 total data-points) for testing the prediction accuracy of their GC

model for aqueous solubility. It is to be noted that GC methods, unlike QSPR type property models where the challenge is the selection of information rich but minimum number of descriptors thereby forming the pattern matrix, the model structure is predefined which contains fixed set of groups (model parameters). Like for most data-driven statistical models, the primary issue with QPSR model development is model discrimination problem (can be defined as model structure identification), which corresponds to identifying a minimum number of descriptors using several methods such as machine-learning algorithms (principal component analysis, support vector machine, etc) such that a valid model can be obtained with minimum cross validation error. In GC based property model development, the primary challenge is the identifiability of model parameters (read as parameter estimation problem) since model structure is given by group-definitions. Hence, in this regard, model validation focus is different for QPSR and GC based models; in particular, the primary concern in GC based models is having a representative data set from which the model parameters can be estimated. This among others means that one cannot randomly divide the total data-set into the training versus validation for GC model development as is done for QSPR models. For model quality evaluation, in GC based models one could instead evaluate the uncertainty (such as 95% confidence interval) on the estimated parameters as well as the uncertainty in the resulting prediction error of the model, which is commonly done in model identification problems (see section 3.2.4 of chapter 3). In addition, through a systematic analysis (see section 3.2.1 of chapter 3), one could identify the minimum data-set needed to estimate the model parameters and in this way, the remaining data-points not used for parameter estimation could be used to evaluate the prediction accuracy of the developed GC model.

2.4.2 Selection of appropriate form of property model function

The GC methods are based on the semi-empiricism and hence, it is imperative to select a suitable form of the property model based on the knowledge of the pure component property and the property data. For example, a plot of normal melting point (T_m) of various classes of pure components versus their increasing carbon number (see Figure 2.3) shows that T_m increases asymptotically with the carbon number suggesting that the appropriate form of property model function is of the exponential form. This trend is also true for the critical temperature (T_c) and normal boiling point (T_b) of pure components. However, in the literature many authors (for example, Joback and Reid, 1987; Simamora and Yalkowsky, 1994; Jain et al., 2004) have employed linear form of GC model function for modeling T_m of pure components. Constantinou and Gani (1994) have chosen the exponential form of model function for modeling of T_c, T_b, and T_m.

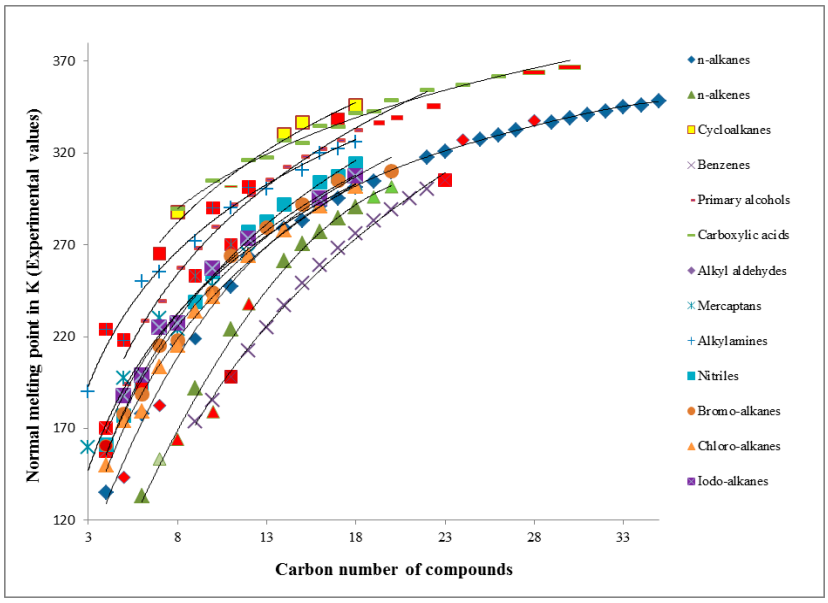


Figure 2.3 Plot of experimental values of T_m versus carbon number of pure components

2.4.3 Uncertainty analysis for evaluating reliability of predicted property values

For assessing the quality and reliability of the predicted property values, it is important to quantify and analyze the uncertainties of the predicted property values. The uncertainty of property data/model predictions (which is a range of values believed to include the "true value" with a certain probability) is a measure of the quality of the measured data/predicted property values. The uncertainties of predicted property values arise due to the approximate nature of the selected property model as well as due to uncertainties of experimentally measured property values used in the parameter regression. The uncertainty analysis of property models can be performed using the frequentist approach (for example, maximum-likelihood estimation) or by using Bayesian approach such as, Markov Chain Monte Carlo algorithms (Gelman et al., 2004). Although, the issue of uncertainty analysis of property models is important and critical, very few studies addressing this issue have been reported in literature. Maranas (1997) followed a covariance matrix based approach to quantify the uncertainties of predicted values that were needed in the computer-aided optimal design of polymers. The author used a multi-linear regression to compute the unbiased estimators (means) and the sample estimators of the variance-covariance matrix of the vector of uncertain group contributions. Frenkel et al. (2005) discussed a new concept of dynamic data evaluation known as ThermoData Engine (TDE) developed at Thermodynamics Research

Center of National Institute of Standards and Technology (NIST/TRC®). Uncertainties of predicted property values, U, are calculated by the following Eq. (2.7).

$$U = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} \cdot \frac{\partial X}{\partial \boldsymbol{P}_i} \cdot \frac{\partial X}{\partial \boldsymbol{P}_j} \right)^{1/2} \tag{2.7}$$

Where, C_{ij} are elements of the covariance matrix, n are the number of model parameters, and $\delta X/\delta \boldsymbol{P}$ is the first derivative of the property X with respect to model parameters, $\boldsymbol{P}$. Recently, Hajipour and Satyro (2011) discussed the uncertainty analysis of property models for critical constants and acentric factor using a covariance matrix based approach. These authors have also considered the uncertainties of experimentally measured property data collected from TDE as weighing factors for data-points (better quality means higher weight and more reliability) in weighted least square regression analysis.

The information of uncertainties of property estimates helps design engineers to evaluate the impact of these uncertainties on the quality of product-process design and take corrective measures, for example, select a property prediction model with "acceptable" uncertainties.

2.4.4 Themodynamic consistency and predictive power of GC models

While developing GC based models for the estimation of pure component properties, it is necessary to perform consistency tests and extrapolation analysis so that the predicted property values are qualitatively correct and safe when properties are estimated outside the proven range of property models. Figure 2.4 shows a plot of ratio of T_c to T_b of n-alkanes predicted using the Joback and Reid method (1987) versus increasing carbon number of n-alkanes up to 100. This plot shows inconsistency with respect to physical principle, that is, the ratio T_c/T_b for a pure component must be always greater than unity. The main reason for this inconsistency is that Joback and Reid (1987) selected an inappropriate form of the model function (binomial equation as shown in Eq. 2.8, instead of exponential equation) for modeling T_c of pure components.

$$T_c = T_b \left[0.584 + 0.965 \sum_i N_i C_i - \left(\sum_i N_i C_i \right)^2 \right]^{-1} \tag{2.8}$$

It can be seen from Figure 2.4 that, the ratio T_c/T_b predicted by Constantinou and Gani (1994) method is always greater than unity.

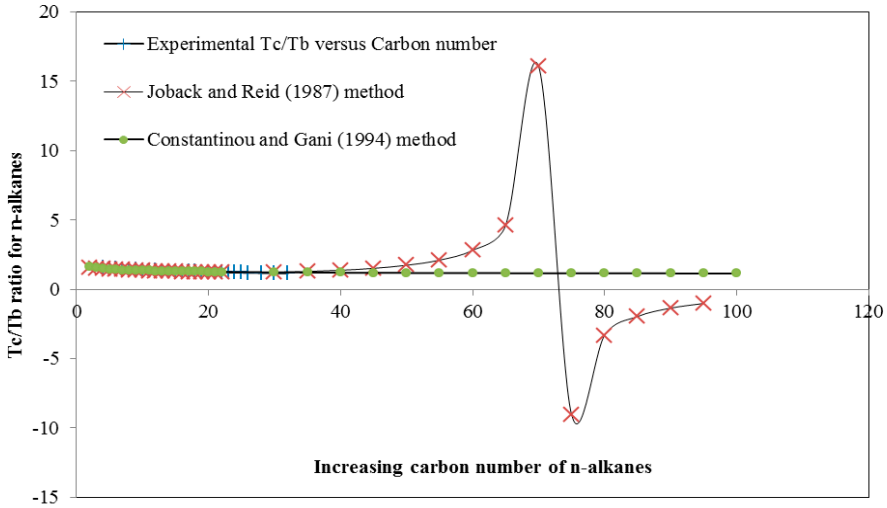


Figure 2.4 Plot of T_c/T_b for *n*-alkanes versus their increasing carbon number

2.4.5 Sensitivity of process design due to uncertainties of property estimates

In chemical process design, the property data/models belong to the constitutive equations that are part of the process model. Hence, the need of property data/models depends on the process/unit operations under consideration, and also on the type (and complexity) of the process model employed. Larsen (1986) pointed out that, in process design, some properties are more important than others, and their importance varies from one process to the other. According to the author, the consequences of data errors can be either catastrophic or incremental. Catastrophic consequences involve a major discontinuity in the process operation including failure of process equipment and plant shut down. Incremental consequences, on the other hand, may not force the plant to be shut down but may result in lower throughout, lower product purity, and higher capital and operating cost. To indicate the effects of the errors of the physical property data, the author has discussed many practical examples. For example, a 20% error in liquid density value may result in a 16% error in equipments size or cost, but a 20% error in diffusivity value may result in only 4% error in equipment size or cost. However, densities are usually known to better accuracy than 20% error, so density errors may be insignificant, while errors in diffusivity may be a factor of two or greater, so that such errors may become very important.

Brignole et al. (1985) discussed a simple algorithm for performing sensitivity and operability analysis of separation processes. The analytical expression developed for calculating number of stages for the rectifying section (N_R) and for the stripping section (N_S) of the distillation column was used study the sensitivity of N_R. N_S, reflux ratio, and reboiler duty due to prediction errors of relative volatilities using a local sensitivity analysis approach. It was

shown that as the system relative volatility approaches to unity, more accurate predictions of thermodynamic properties are required. Dohrn and Pfohl (2002) also showed that, for difficult separations, an underestimation of the correct value of α =1.1 by 5% leads to an error in the calculated required distillation column height by more than 100%. Macchieto (1986) addressed the problem of establishing which properties and parameters, if any, are critically important in a given process. The author employed a local sensitivity analysis to analyze the sensitivities of process variables with respect to constant parameters and functions of correlations of different property models. Whiting et al. (1993) addressed the importance of consideration of uncertainties of both the property data and property model parameters in process design through several case studies. These authors employed Monte Carlo simulation approach with Latin-Hypercube Sampling (LHS) technique to quantify effects of uncertainties in property data/model parameters on the calculated performance of unit operations such as super-fractionator and flash process. These case studies highlight the importance of performing sensitivity analysis of process design subject to uncertainties in the property data/models so that the effects of these uncertainties on the quality of process design can be evaluated. The results of this analysis can be further evaluated to take necessary actions (for example, select appropriate safety factors) to improve the final design.

2.5 Summary

GC based property prediction models are frequently employed to obtain necessary thermo-physical, transport-related, and environmental-related pure component properties in the design and analysis of chemical products and processes. Significant developments and improvements have been made in the field of property modeling including, development of more accurate property models, increased application range, and improved predictive power of GC models. However, the current literature review indicates that there are challenges and issues related to property modeling that need to be carefully addressed which is presented in the following chapters.

Chapter 3. Methodology

3.1 Basis for Selecting Marrero and Gani (MG) Method

The MG method allows estimation of properties of pure components based exclusively on the molecular structure of the component and exhibits a good accuracy and a wide range of applicability covering chemical, biochemical and environmental-related components. The MG method takes into account the interactions among different groups (for example, complex interactions among groups in heterocyclic species) to increase the accuracy and reliability of predicted property values. As discussed in section 2.1 of chapter 2, the MG method employs property estimation at three levels. The first-level estimation provides an initial approximation using only first-order group-contributions that is improved at the second-level using second-order group-contributions which is further refined at the third-level using third-order group- contributions. The main objective of this multi-level estimation scheme is to improve the accuracy, reliability, and application range of property models for a number of pure component properties. Therefore, in this work, we have selected the MG method based GC models for developing property models for the estimation of pure component properties. For property modeling with a CI method, the models proposed by Gani et al. (2005) have been considered. These CI models are employed together with the MG method for the purpose of creating missing groups and predicting their missing group contributions. As discussed in chapter 2, the property prediction model employing the MG method has following the form:

$$f(X) = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k E_k O_k \tag{3.1}$$

In MG method, the molecular structure of each pure component is decomposed into the first-order, second-order, and third-order MG groups. The MG method includes 220 first-order groups, 130 second-order groups, and 74 third-order groups to represent the molecular structure of pure components. In the development of CI method based property models, the molecule is decomposed into atoms present in that molecule and these atoms represent as the structural parameters for the CI method based property models. For the decomposition of molecular structure of pure components into MG groups and atoms, the ConvertSmiles program available as a toolbox within the ICAS® software is used. The ConvertSmiles program takes SMILES (SMILES is an acronym for Simplified Molecular Input Line Entry System.) of pure components as the input and then decomposes their molecular structures into the first-order, second-order, and third-order MG groups and atoms. The zeroth-order and first-order connectivity indices of pure components that are required in the development of CI method based property models can also be obtained through ConvertSmiles program.

For the determination of contributions, C_i, D_j, and E_k, Marrero and Gani (2001) suggested a multi-level estimation approach.

- *Level 1*: In this level, the constants *w* and *z* are assigned zero values because only contributions of the first-order groups are estimated, that is, the first-order groups, C_i and the additional adjustable parameters of the model.

$$f(X) = \sum_i N_i C_i \tag{3.2}$$

- *Level 2*: In this level, the constants *w* and *z* are assigned unity and zero values, respectively, because only first and second-order groups are considered. The regression is performed (by keeping fixed the C_i and the adjustable parameters obtained from level 1) to determine the contributions of the second-order groups, D_j.

$$f(X) = \sum_i N_i C_i + \sum_j M_j D_j \tag{3.3}$$

- *Level 3*: In this level, both *w* and *z* are set to unity and regression is performed (by keeping fixed the obtained C_i, D_j, and the adjustable parameters obtained from level 1 and 2) to determine the contributions of the third-order groups, E_k.

$$f(X) = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k E_k O_k \tag{3.4}$$

In this work, a new approach for estimating the contributions, C_i, D_j, and E_k based on the simultaneous regression method is developed in which regression is performed by considering all the terms of Eq. (3.1) to obtain contributions of first-order, second-order, and third-order MG groups in a single step.

3.2 Methodology for Performing Property Modeling and Analysis

The challenges and issues related to property modeling as discussed in section 2.4 of chapter 2 are addressed through a systematic methodology for performing property modeling and analysis of pure component properties and is presented in following sub-sections 3.2.1 to 3.2.6. The application of this methodology is illustrated in chapter 4 through development and analysis of GC^+ models for estimation of pure component properties and in chapter 6 though three case studies namely, design of an extractive distillation process, design of a short-path evaporator, and design of a de-acidification system of vegetable oil deodorization process.

3.2.1 Selection of quantity of property data for parameter regression

In this work, we have considered all of the available experimental data-points while modeling of pure component properties for the reason discussed in section 2.4 of chapter 2. However, in some cases, while developing a GC model, it is preferred to retain some data-points from the total data-set to test the reliability of predicted property values. In such cases, the selection of a minimum data-set for the parameter regression becomes very important. One important criterion for selecting the minimum data-set is that the information available in the original data-set should also be available in the selected minimum data-set such that the model parameters can be identified properly. That is, the number of GC model parameters estimated by regressing the total data-set must be equal to the number of GC model parameters estimated using the selected minimum data-set. For linear systems, the minimum data-set for the regression can be identified from the total data-set using simple rules of linear algebra. To this end, the theorem of Rouché–Capelli (Shafarevich and Remizov, 2013) can be used which states that the system A$\mathbf{x}$ = $\mathbf{b}$ admits unique solutions if and only if,

$$\rho(\mathrm{A}) = \rho(\mathrm{A}|\mathbf{b}) \tag{3.5}$$

Where A is the coefficient matrix, $\mathbf{x}$ is the solution vector and $\mathbf{b}$ is a column vector. The matrix A|$\mathbf{b}$ is the augmented matrix containing A and $\mathbf{b}$. For the problem at hand, A is the occurrence matrix (of size $m \times n$) containing occurrences of all groups included in a GC-method under consideration, $\mathbf{b}$ is the vector (of size $m \times 1$) containing experimentally measured values of compounds included in the data-set, and $\mathbf{x}$ is the solution vector (of size $n \times 1$) containing model parameters (that is, group-contributions) whose values are to be determined by solving A$\mathbf{x}$ = $\mathbf{b}$. The m and n corresponds to the number of data-points in the data-set and number of unknown GC model parameters respectively. In Eq. (3.5), ρ(A) is the rank of matrix A and ρ(A|$\mathbf{b}$) is the rank of augmented matrix containing occurrences of all groups and experimentally measured values represented by $\mathbf{b}$. The degree of freedom, ν, is defined as (n - ρ(A)) and by definition the solutions $\mathbf{x}$ exists if and only if ν is zero. In the context of finding a minimum data-set, we select p number of data-points from the total data-set and solve new system of equations, $\mathrm{A}^*\mathbf{x} = \mathbf{b}^*$, such that,

$$\rho\left(\mathrm{A}^*\right) = \rho\left(\mathrm{A}^*\middle|\mathbf{b}^*\right) \tag{3.6}$$

Where, $\rho(\mathrm{A}^*)$ is the rank of matrix A^* which is sub-matrix of A and $\rho(\mathrm{A}^*|\mathbf{b}^*)$ is the rank of augmented matrix $\mathrm{A}^*|\mathbf{b}^*$ which is sub-matrix of A|$\mathbf{b}$. An additional and most important criterion is that the degree of freedom should be zero which implies that the number of unknown model parameters identifiable from the original total data-set is still equal to number of unknown model parameters identifiable from the selected minimum data-set. This

criterion ensures that the application range of the GC model is not affected even when the minimum data-set for parameter regression is extracted from the original data-set. Note that in Eq. (3.6), the effect of measurement error on $\boldsymbol{b}$ has not been considered here. To illustrate the application this approach let us consider a data-set of enthalpy of formation, $\Delta_f H^o_{gas}$, containing 4 *n*-alkanes and 4 *n*-alkenes as shown in Table 3.1.

Table 3.1 Occurrence matrix and experimental data of $\Delta_f H^o_{gas}$ for *n*-alkanes and *n*-alkenes

Compounds in the data-set	Groups and their occurrence matrix (A)			Experimental values (**b**) of $\Delta_f H^o_{gas}$ in kJ/mol
	CH_3 **(x_1)**	CH_2 **(x_2)**	CH_2=CH **(x_3)**	
Pentane	2	3	0	-146.76
Heptane	2	5	0	-187.65
Decane	2	8	0	-249.46
Dodecane	2	10	0	-290.72
1-Pentene	1	2	1	-21.30
1-Hexene	1	3	1	-42.00
1-Octene	1	5	1	-83.60
1-Decene	1	7	1	-124.70

For this example, ρ(A) is 3 and ρ(A|**b**) is 4. In the given example, the three model parameters (that is, groups) denoted as $\mathbf{x_1}$, $\mathbf{x_2}$ and $\mathbf{x_3}$ are to be estimated. The rank of augmented matrix is 4 which indicate that there are four independent equations, and hence more than sufficient to estimate 3 unknowns parameters. This information suggests that we can remove some of the data-points from the total data-set which for example can be used for the validation purpose. By performing an analysis of removal of a particular data-point from the data-set and simultaneously checking the ranks of the corresponding coefficient matrix (ρ (A^*)) and the rank of the augmented matrix ($\rho(A^*|\mathbf{b}^*)$) one can obtain minimum number of data-points required to estimate the unknown GC model parameters namely, CH_3 ($\mathbf{x_1}$), CH_2 ($\mathbf{x_2}$), and CH_2=CH ($\mathbf{x_3}$). For the above example, we need to choose 3 components for the regression purpose. These are very minimum number of data-points needed for the regression (estimation of *n* unknown model parameters requires *n* data-points). However, the model developer needs to retain sufficient number of data-points (over the true minimum data-points) for the regression purpose so that accurate and reliable predictions are achieved. For this example, let us consider the minimum data-set for parameter regression consists of 3 *n*-alkanes (pentane, heptane, and decane) and 3 *n*-alkenes (1-Pentene, 1-Hexene, and 1-Octene). After solving Eq. (3.6) we obtain group-contributions as CH_3 = -42.36; CH_2 = -20.60; and CH_2=CH = 62.08. The average absolute error obtained from the parameter regression is 0.17 kJ/mol. The extra available data-points namely, Decane, and 1-Decene can be used for the

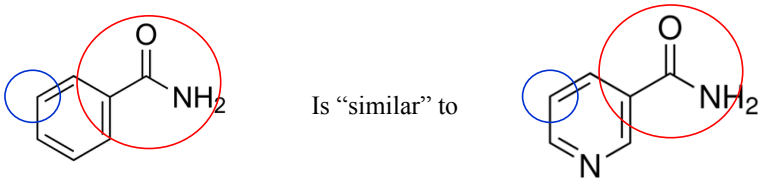
model validation purpose. Using these regressed group-contributions, we estimate $\Delta_f H^o_{gas}$ of Decane as -290.72 kJ/mol (with an error of 0.06 kJ/mol) and for 1-Decene as -124.52 kJ/mol (with an error of 0.17 kJ/mol). The average absolute error for these two validation pure components is therefore 0.12 kJ/mol. It can be seen that the group-contributions regressed using the minimum data-set are able to provide accurate predictions of $\Delta_f H^o_{gas}$ for validation components.

3.2.2 Selection of suitable property prediction model function

The most appropriate form of the model function, *f(x)*, can be derived by analyzing the behavior of properties of various classes of pure components as a function of their carbon number. When the experimentally measured data of pure components belonging to different classes such as *n*-alkanes, primary alcohols, ketones, ethers etc. are plotted against their carbon number, a trend which can be a linear (as in the case of critical volume), exponential (as in the case of normal boiling point) or a polynomial can be observed. This observed trend is then taken as a basis for selecting a suitable form of *f(x)*.

3.2.3 'Molecular structural similarity criteria' for improvement of property models

To improve the performance of developed GC model and to achieve the required level of prediction accuracy, sometimes it is necessary to include new structural parameters (that is, groups) in the GC model to provide more structural information to pure components having large prediction errors. The 'molecular structure similarity criteria' based approach can be used as a basis for defining new structural parameters of the GC model. In this approach, the molecular structure of one pure component is compared with the molecular structures of other pure components in the data-set to identify a set of pure components that are "similar" in nature. Here, "similar" components mean components having one or more functional first-order groups in common. For example, the component Benzamide is "similar" to Nicotinamide in that, the first-order MG groups, aCH and aC-$CONH_2$, are common in the molecular structure of both these components (see Figure 3.1). Detailed description and application of 'molecular structure similarity criteria' based approach is further discussed in chapter 4.



Benzamide Nicotinamide

Figure 3.1 Application of 'molecular structural similarity criteria'

3.2.4 Parameter estimation and uncertainty analysis using maximum-likelihood estimation theory

For notational convenience, the property prediction model is represented by f and the model parameters (for example, group contributions or atom contributions) by $\boldsymbol{P}$. It is to be noted that the unknown model parameters, $\boldsymbol{P}$, can be approximated by its estimators $\boldsymbol{P}^*$ by using a proper estimation theory. Non-linear least squares method is used which is a special case of maximum likely-hood estimation theory in which the minimization of a cost function, $S(\boldsymbol{P}^*)$, which is the sum of the squares of the difference between the experimental value, X^{exp} and evaluated property value, X^{pred} yields the values of unknown parameters $\boldsymbol{P}^*$. Note that $\boldsymbol{P}^*$ is the set of model parameter values obtained at the minimum value of the cost function value.

$$S\left(\boldsymbol{P}^*\right) = min \sum_{j=1}^{N}\left(X_j^{exp} - X_j^{pred}\right)^2 \tag{3.7}$$

The subscript j indicates the pure component evaluated and N is the total number of pure components included in the evaluation.

To infer about the quality of parameter estimation, covariance matrix can be computed. For non-linear least-squares, the covariance matrix of the estimated model parameters can be approximated from a linear approximation of the model (Seber and Wild, 1989). The linear approximation of the covariance matrix, $COV(\boldsymbol{P}^*)$, for the estimated parameters is given by:

$$COV\left(\boldsymbol{P}^*\right) = \frac{SSE}{\nu}\left(J\left(\boldsymbol{P}^*\right)^T J\left(\boldsymbol{P}^*\right)\right)^{-1} \tag{3.8}$$

Where, SSE is the minimum sum of squared errors obtained from the least-squares parameter estimation method, v is the degree of freedom (that is, the total number of measurements, m minus the number of unknown parameters, n). The Jacobian matrix $J(\boldsymbol{P}^*)$ calculated using $\partial f/\partial \boldsymbol{P}^*$represents the local sensitivity of the property model f to variations in the estimated parameter values $\boldsymbol{P}^*$. The covariance matrix computed using Eq. (3.8) is used for assessing the quality of parameter estimation. The diagonal elements of this matrix are the variances of the errors of the parameter estimates and the off-diagonal elements are the covariances between the parameter estimation errors.

For linear least squares, the covariance matrix of the estimated model parameters is given by:

$$COV\left(\boldsymbol{P}^*\right) = \frac{SSE}{\nu}\left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1} \tag{3.9}$$

For the GC based property model with linear form of $f(X)$, $\boldsymbol{A}$ is the matrix containing frequencies (or occurrences) of groups used to represent the molecular structure of pure

components in the data-set for the regression. For the CI method based model with linear form of $f(X)$, $\boldsymbol{A}$ is the matrix containing frequencies of atoms and zeroth-order and first-order connectivity index for each chemical included in the data-set. The error on the estimated property values can be calculated via linear error propagation (Seber and Wild, 1989) as follows:

$$COV\left(X^{pred}\right)=\left(J\left(\boldsymbol{P}^{*}\right)\ COV\left(\boldsymbol{P}^{*}\right)\ J\left(\boldsymbol{P}^{*}\right)^{T}\right) \tag{3.10}$$

The confidence interval of the parameters, $\boldsymbol{P}^{*}$, at α_t significance level is given as:

$$\boldsymbol{P}^{*}_{1-\alpha_t}=\boldsymbol{P}^{*}\pm\sqrt{\text{diag}\left(COV\left(\boldsymbol{P}^{*}\right)\right)}\ .\ t\left(\nu,\ \alpha_t/2\right) \tag{3.11}$$

In Eq. (3.11), t (v, α_t /2) is the t-distribution value corresponding to the $\alpha_t/2$ percentile (α_t is usually a value of 0.05) and diag(COV($\boldsymbol{P}^{*}$)) represents the diagonal elements of COV($\boldsymbol{P}^{*}$). The t-distribution value is obtained from the probability distribution function of Students t-distribution, (Abramowitz and Stegun, 1972), $P_r(t, \nu)$, and is given as:

$$0=\sqrt{\nu}B\left(\frac{1}{2},\frac{\nu}{2}\right)^{-1}\int_{-t}^{t}\left(1+x^2/\nu\right)^{-\frac{1}{2}(\nu+1)}dx\ -\ P_r\left(t,\nu\right) \tag{3.12}$$

Where $x=\nu/\nu+t^2$ and B (1/2, ν/2) is the beta function. For 95% confidence interval calculation, the value of $P_r(t, \nu)$ is 0.95. The t-distribution value can also be obtained using the "tinv function" available in Matlab or from any of the number of reported statistical tables. The confidence interval of the predicted property value, X^{pred}, at α_t significance level is given as:

$$X^{pred}_{1-\alpha_t}=X^{pred}\ \pm\sqrt{\text{diag}\left(J\left(\boldsymbol{P}^{*}\right)COV\left(\boldsymbol{P}^{*}\right)J\left(\boldsymbol{P}^{*}\right)^{T}\right)}\ .\ t\left(\nu,\alpha_t/2\right) \tag{3.13}$$

It is to be noted that the uncertainties of the property data arise mainly due to accuracy and precision of measurement instruments used, method of measurement, and purity of samples considered in the analysis, among others. The model prediction error reported as 95% confidence interval on calculated properties is a statistical concept associated with statistical framework used for parameter estimation. In this work the maximum likelihood theory as summarized in Eqs (3.7) − (3.13) is employed, which aims to propagate the residuals (that is, the difference between the property data and the calculated property values using the model) obtained after parameter estimation as first errors on model parameters (covariance matrix of

estimated parameters) and then errors on the model predictions using linear error propagation method (Seber and Wilde, 1989). The $X_{1-\alpha_t}^{pred}$ calculated from Eq. (3.13) provides the confidence interval of the predicted property value at a specified confidence level and it can be used to assess the reliability of the prediction (when experimental data is available for the property). If the experimental value of the property is within the calculated confidence interval, then it can be concluded that the predicted property value (and hence the prediction method) is reliable. When no experimental data is available, the calculated confidence interval provides a measure of the likely prediction error (uncertainty) of the predicted property value. In this work, for evaluating the performance of developed property models on a global basis, following statistical performance indicators (Marrero and Gani, 2001) are used.

Standard Deviation (SD): This parameter measures the spread of the data about its mean value μ and is given by:

$$\mathrm{SD} = \sqrt{\sum_j \left(X_j^{exp} - X_j^{pred} \right)^2 / N} \tag{3.14}$$

Average Absolute Error (AAE): This is the measure of deviation of predicted property values from the experimentally measured property values and is given by:

$$\mathrm{AAE} = \frac{1}{N} \sum_j \left| X_j^{exp} - X_j^{pred} \right| \tag{3.15}$$

Average Relative Error (ARE): This provides an average of relative error calculated with respect to the experimentally measured property values and is given by:

$$\mathrm{ARE} = \frac{1}{N} \sum_j \left| \left(X_j^{exp} - X_j^{pred} \right) / X_j^{exp} \right| \times 100 \tag{3.16}$$

Coefficient of Determination (R^2): This parameter provides information about the goodness of model fit. An R^2 close to 1.0 indicates that the experimental data used in the regression have been fitted to a good accuracy. It is calculated using:

$$\mathrm{R}^2 = 1 - \left[\sum_j \left(X_j^{exp} - X_j^{pred} \right)^2 / \sum_j \left(X_j^{exp} - \mu \right)^2 \right] \tag{3.17}$$

3.2.5 Tests for thermodynamic consistency and predictive power

To verify the thermodynamic consistency and predictive power of developed property models, the tests discussed by Tsonopoulos (1987) and Tsonopoulos and Tan (1993) are employed in this work. These include: (i) check whether limiting values of critical pressure,

critical density are approached; (ii) T_c of pure component should be always greater than T_b; and (iii) the critical pressure of pure components decrease with increasing carbon number of pure components. The upper limit of 0.291 (Poling et al., 2000) of critical compressibility factor of pure components having T_c greater than 100 K is also tested.

3.2.6 Sensitivity analysis of process design due to uncertainties of property estimates

In this section, a step by step description of the developed method for performing sensitivity analysis of process design due to uncertainties of property estimates is presented (see Figure 3.2). With the results of sensitivity analysis of process design due to uncertainties of property estimates, it is possible to answer following three questions: (i) what are the effects of uncertainties of property estimates on the final design parameters? (ii) which properties are the most sensitive from a process design point-of-view? and (iii) what is the acceptable levels of accuracy for different thermo-physical property prediction models employed in the design?

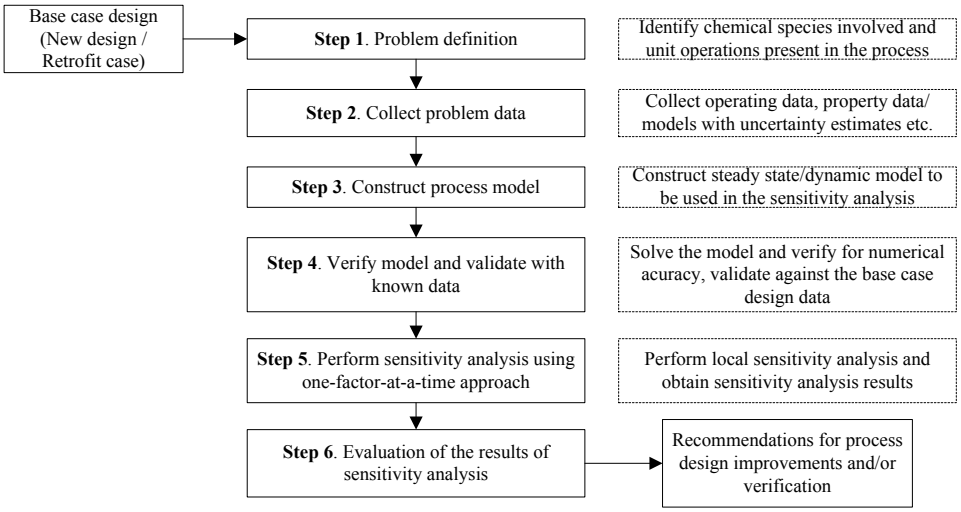


Figure 3.2 Methodology for performing the sensitivity analysis for process design

Step 1. Problem definition

In this step, the process (base case) for which sensitivity analysis is to be performed is defined. Information such as the chemical species involved, the unit operations present in the process is also defined.

Step 2. Collect problem data

The problem data is collected in this step. This includes operating data of the process, desired specifications, kinetics data (if reactions are involved), data related to economic evaluation,

equipment data, constraints (operational, environmental etc.), physical and thermodynamic data and models of concerned properties together with tools to quantify uncertainties of predicted property values.

Step 3. Construct process model

The objective of this step is to construct a model for the process under study so that the effects of change in property estimates can be analyzed. The model may be steady state or dynamic, depending on the problem definition. Note that the property models belong to the constitutive equations that are part of the process model. This task can be achieved in a suitable modeling environment such as ICAS-MoT® software or in a commercial process simulator such as PRO/II®.

Step 4. Verify model and validate with known data

The process model is solved. The obtained simulation results are verified for numerical accuracy and validated against base case design data.

Step 5. Perform sensitivity analysis using one-factor-at-a-time approach

Local sensitivity analysis such as one-factor-at-a-time approach is employed. The local sensitivity, J, as shown in Eq. (3.18), is the sensitivity of the model output (y) when only one input factor, Θ, is perturbed at a time keeping all other input factors fixed at their base case design values (Sin et al., 2010).

$$S = \frac{\partial y}{\partial \theta} \tag{3.18}$$

The obtained sensitivity analysis results are represented in the form of plots showing relationships between design variables and the involved properties. The y-axis represents the deviation (expressed as relative error in %) of the design variables from its base case design value and the x-axis represents the expected prediction error (expressed in %) in the properties.

Step 6. Evaluation of the results of sensitivity analysis

The objective of this step is to analyze the relationships between design variables and properties and evaluate the impact of uncertainties of predicted properties on the design; list properties based on their sensitivity; and provide an acceptable level of accuracy for different property models. This involves: (i) examining the slope of plot of design variable versus prediction error in the property and identifying the most sensitive properties (a large slope implies that sensitivity of the design variable is high and the uncertainties of property estimates are significant); (ii) analyzing the possible consequences (such as, effect on capital and operating cost, inefficient operation, feasibility of the operation of a given process) of

property errors on the process design; and (iii) establishing an acceptable level of accuracy for property models based on the maximum allowable deviation (as specified by the user) in the design variable from its base case design value. The analysis carried out in step 6 can be used as an input to a process design tool for improvement in the design.

3.3 Summary

In this chapter, the developed methods that are necessary in the development and analysis of property models for pure component properties are presented. These methods assists the model developer to select the quantity of the data for the parameter regression; select suitable form of the model function; achieve the required level of prediction accuracy through 'molecular structural similarity criteria'; quantify uncertainties (for example, 95% confidence intervals) of predicted property values; and address issues such as thermodynamic consistency and predictive power of property models. In chapter 4, the application of these methods is highlighted through the development and analysis of GC^+ based property models for thermo-physical, transport-related, and environmental-related properties of pure components. The application of the developed method involving "molecular structure similarity criteria" approach is highlighted through performance improvement of GC models for enthalpy of formation, enthalpy of fusion, and critical temperature of pure components. Finally, in chapter 6, the application of method for performing sensitivity analysis of process design due to uncertainties of property estimates is illustrated using three case studies.

Chapter 4. Results: Model Development and Verification

4.1 Development of GC$^+$ Models for Thermo-Physical and Transport-Related Properties

The following 21 thermo-physical and transport properties of pure components are considered for the property modeling: normal boiling point (T_b), critical temperature (T_c), critical pressure (P_c), critical volume (V_c), normal melting point (T_m), standard Gibbs energy of formation ($\Delta_f G^o$), standard enthalpy of formation ($\Delta_f H^o_{gas}$), normal enthalpy of fusion ($\Delta_{fus}H$), enthalpy of vaporization at 298 K ($\Delta_{vap}H°$), enthalpy of vaporization at the normal boiling point ($\Delta_{vap}H$), entropy of vaporization at the normal boiling point ($\Delta_{vap}S$), flash point (F_p), auto ignition temperature (Ait), Hansen solubility parameters (δ_D, δ_P, and δ_H), Hildebrand solubility parameter (δ), octanol/water partition coefficient ($Logk_{ow}$), acentric factor (ω), liquid molar volume at 298 K (V_m), liquid viscosity (μ), liquid surface tension (σ), and liquid thermal conductivity (k). In this work, two pure component properties namely, ω and V_m which were modelled earlier by Constantinou and Gani (1995) are modelled using the MG method in order to provide improved property estimations for a wide range of pure components. The results are presented for the following pure component property prediction models.

- MG method based property models analyzed using step-wise regression method
- MG method based property models analyzed using simultaneous regression method
- CI method based property models

4.1.1 Database

Since the publication of the original work on the property modeling by Gani and co-workers, significant amount of new experimental data of pure components (especially poly-functional, polycyclic, and complex components) have been added to the CAPEC database (Neilsen et al., 2001) developed and extended at CAPEC, DTU. The extended CAPEC database containing large experimental data-sets of pure components of various classes is used with the aim of providing improved property estimation of pure component properties. In addition, the newly available experimental data helps in increasing the application range of the analysed property models through estimation of model parameters (group/atom contributions) whose values were not determined and published in the previous works due to the lack of necessary experimental data. The details of data-set of each property in terms of number of components belonging to various classes are given in Table 4.1.

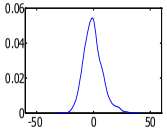
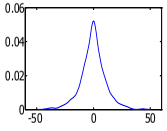
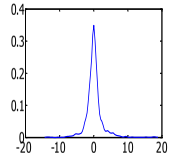
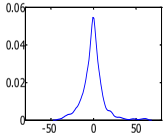
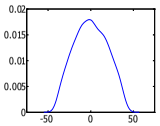
4.1.2 Model performance statistics

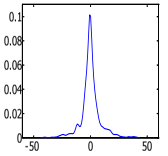
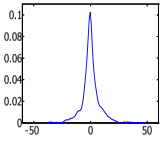
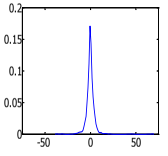
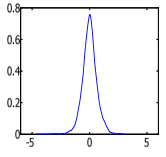
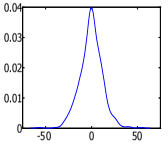
The model performance statistics for the property models analysed using the step-wise regression method and the CI method is summarized in Table 4.2 and Table 4.3 respectively. The model performance statistics for properties analysed using simultaneous regression method are given in Appendix A (see Table A1). In Table 4.2 and Table 4.3, N is the number of data-points considered in the regression and v is the degree of freedom for each property and is obtained by subtracting number of estimated model parameters from N. $P_{rc(\pm 1\%)}$, $P_{rc(\pm 5\%)}$, and $P_{rc(\pm 10\%)}$ represents the percentage of the experimental data-points (N) found within ± 1%, ± 5%, and ± 10% relative error range respectively. AE_{max} is the maximum absolute error and RE_{max} is the maximum relative error obtained in the regression analysis. For property models analysed using step-wise regression method, the results for R^2, SD, AAE and ARE have been obtained after third-level estimation and hence they represent the global results. The property models analysed using simultaneous regression method have shown some improvement in the performance as compared to the performance of property models analysed using step-wise regression method. The residuals (X^{exp}-X^{pred}) for data-points considered in the regression are plotted in the form of residual distribution plots and are included in Tables 4.2 and 4.3. From these distribution plots it can be seen that the residuals obtained from the models, except for T_m, follow a normal distribution curve with mean zero suggesting a good fit of the experimental data used in the regression and there is no apparent bias in the predicted property values as well as the assumption that the residuals follow a normal distribution behind the maximum likelihood estimation approach is valid. The use of the CI method for creating the missing groups and predicting their contributions through the regressed contributions of connectivity indices as suggested by Gani et al., 2005 has been implemented. This allows one to make predictions for a number of properties for which neither experimental data nor the GC model parameters are available. The CI method based property models for T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H°$, $\Delta_{vap}H$, ω, and V_m are included in this work.

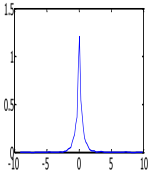
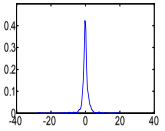
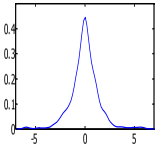
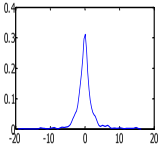
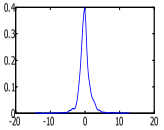
Table 4.1 Class-wise description of the CAPEC database used in the parameter regression

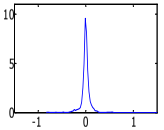
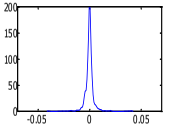
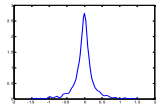
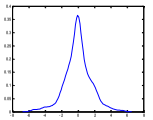
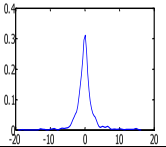
Class of pure components	T_b	T_c	P_c	V_c	T_m	$\Delta_f G^o$	$\Delta_f H^o$	$\Delta_{fus} H$	$LogK_{ow}$	F_p	δ_D	δ_P	δ_H	$\Delta_v H°$	$\Delta_{vap} H$	$\Delta_{vap} S$	δ	Ait	ω	V_m	μ	σ
Hydrocarbons	662	277	288	281	492	279	272	270	233	168	73	57	64	189	112	112	386	191	430	324	102	86
Oxygenated	1187	297	314	280	1493	266	245	237	1500	229	443	444	444	229	185	185	498	243	659	352	163	155
Nitrogenated	369	91	86	76	374	72	149	73	785	42	76	76	75	91	67	67	125	45	140	76	51	45
Chlorinated	202	38	30	29	149	29	28	18	271	21	75	74	74	34	30	30	62	28	65	49	16	12
Fluorinated	46	21	11	8	41	6	5	10	27	-	18	15	12	15	17	17	26	3	41	29	4	3
Brominated	99	8	8	8	89	8	10	11	41	5	29	29	29	23	12	12	20	6	19	15	16	14
Iodinated	30	4	5	5	27	5	5	2	10	-	9	9	9	9	7	7	8	-	8	8	-	-
Phosphorous containing	2	2	-	1	2	-	1	1	2	-	-	-	-	-	-	-	1	-	1	1	-	-
Sulfonated	109	34	33	32	83	31	29	36	65	2	23	23	23	51	39	39	35	2	73	45	3	3
Silicon containing	14	2	2	-	4	-	-	-	6	2	-	-	-	-	-	-	4	4	2	2	3	2
Multifunctional	790	84	75	77	2429	53	138	103	9253	43	291	290	286	64	43	43	219	48	285	155	87	84
Total number of components	3510	858	852	797	5183	749	882	761	12193	512	1037	1017	1016	705	512	512	1384	570	1723	1056	445	404

Table 4.2 Performance of MG method based models for thermo-physical and transport properties analysed using step-wise regression method

Sl. No.	Property	L.H.S. of MG method based property prediction model	**MG group-contribution model** $f(X)=\sum_i N_i C_i + w\sum_j M_j D_j + z\sum_k E_k O_k$											
		$f(X)$	N	v	R^2	Residual distribution plot	P_{rc} (± 1%)	P_{rc} (± 5%)	P_{rc} (± 10%)	SD	AAE	ARE [a]	AE_{max}	RE_{max}
1	Normal boiling point [K]	$\exp\left(\frac{T_b}{T_{bo}}\right)$	3510	3179	0.9980		56.95	99.44	100.0	7.90	6.17	1.44	35.88	11.46
2	Critical temperature [K]	$\exp\left(\frac{T_c}{T_{co}}\right)$	858	607	0.9983		66.03	98.87	100.0	10.77	7.72	1.23	44.34	7.21
3	Critical pressure [bar]	$(P_c - P_{c1})^{-0.5} - P_{c2}$	852	608	0.9690		31.87	80.18	92.34	2.38	1.40	3.90	17.87	25.79
4	Critical volume [cc/mol]	$V_c - V_{c0}$	797	565	0.9956		50.31	93.22	99.00	11.65	7.97	2.05	61.72	21.73
5	Normal melting point [K]	$\exp\left(\frac{T_m}{T_{mo}}\right)$	5183	4803	0.9456		17.60	73.28	93.90	19.16	15.99	5.07	44.43	29.59

6	Gibbs free energy [kJ/mol]	$\Delta_f G^o - G_{f0}$	749	521	0.9984		37.12	72.23	85.31	8.36	5.24	**	40.50	**
7	Enthalpy of formation [kJ/mol]	$\Delta_f H^o{}_{gas} - H_{f0}$	882	649	0.9992		42.86	80.05	90.02	7.74	5.03	**	45.02	**
8	Enthalpy of fusion [kJ/mol]	$\Delta_{fus} H - H_{fus0}$	764	516	0.8324		13.48	32.98	53.40	5.16	2.79	**	68.01	**
9	Octanol/Water partition coefficient	$LogK_{ow} - K_{ow0}$	12193	11817	0.874		5.45	25.44	43.85	0.64	0.48	**	4.35	**
10	Flash point [K]	$F_p - F_{po}$	512	340	0.9671		37.50	89.65	98.44	12.10	8.97	2.8	56.79	12.86

11	Hansen solubility parameter [MPa$^{1/2}$]	Dispersion (δ_D)	1037	769	0.73	 (shown here for δ_D parameter)	39.05	86.89	96.53	1.08	0.60	**	9.52	**
		Polar (δ_P)	1017	754	0.75		9.34	25.66	46.02	2.20	1.81	**	9.68	**
		H_2-bond (δ_H)	1016	754	0.87		12.99	34.06	56.00	2.79	1.28	**	9.24	**
12	Enthalpy of vaporization (298K) [kJ/mol]	$\Delta_{vap}H° - H_{v0}$	705	509	0.9716		40.99	90.07	96.88	2.34	1.29	3.24	26.92	133.9
13	Enthalpy of vaporization (Tb) [kJ/mol]	$\Delta_{vap}H - H_{vb0}$	512	346	0.9606		41.41	90.04	98.05	1.42	0.95	2.66	8.16	22.21
14	Entropy of vaporization (Tb) [J/mol K]	$\Delta_{vap}S - S_{vb0}$	512	346	0.8539		58.20	95.90	98.63	3.0	1.72	1.84	17.98	19.16
15	Hildebrand solubility parameter [MPa$^{1/2}$]	$\delta - \delta_0$	1384	1089	0.8290		34.25	70.45	90.46	1.63	1.08	5.61	12.17	44.49
16	Auto ignition temperature[b] [K]	T_{AiT}	This property is modeled using simultaneous regression method only and hence model performance statistics for this property are provided in Appendix A.											

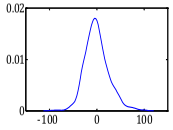
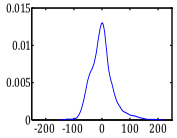
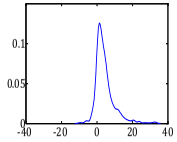
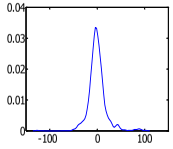
17	Acentric factor	$\exp\left(\frac{\omega}{\omega_a}\right)^{\omega_b} - \omega_c$	1723	1422	0.8921		41.56	56.36	78.58	0.1002	0.0534	11.09	1.68	77.55
18	Liquid molar volume [cc/kmol]	$V_m - V_{m0}$	1056	800	0.9967		69.60	92.90	98.77	0.0045	0.0024	2.03	0.0401	37.05
19	Liquid viscosity [mPa-s]	$\ln(\mu)$	522	339	0.90		14.41	32.21	49.55	1.84	0.72	25.50	1.57	31.6
20	Liquid surface tension [mN/m]	σ	546	373	0.905		32.2	79.84	94.50	1.62	1.39	5.27	5.55	28.2
21	Thermal conductivity [mW/mK]	k	336	198	0.89		51.4	78.9	91.2	10.06	6.34	5.13	12.3	21.7

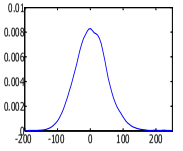
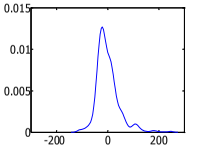
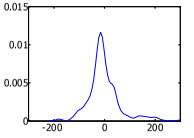
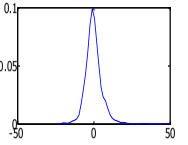
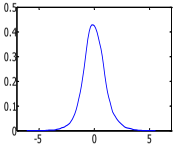
[a] For $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, and $LogK_{ow}$, ARE is not reported since these properties contain both positive and negative values. For Hansen solubility parameters δ_D, δ_P, and δ_H, and liquid viscosity, ARE and REmax is not reported as these properties contain very small experimental values.

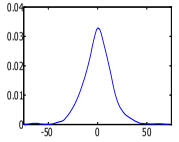
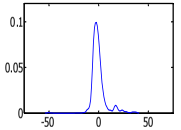
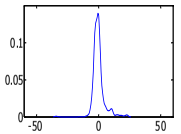
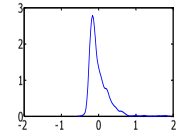
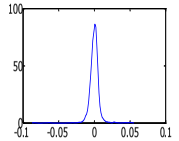
[b] For auto ignition temperature, the right hand side of the MG model is different and it is given as,

$$T_{Ait} = Ait_1\, 10^{-\sum N_i C_{ai} + \sum M_j D_{aj} + \sum E_k O_{ak}} + Ait_2 + \sum N_i C_{bi} + \sum M_j D_{bj} + \sum E_k O_{bk}$$

Table 4.3 Performance of CI method based property models for thermo-physical and transport-related properties

Sl. No.	Property	L.H.S. of CI method based property prediction model	**CI method based property prediction model** $f(X)=\sum_i a_i A_i + b\left({}^{v}\chi^{0}\right)+2c\left({}^{v}\chi^{1}\right)+d$									
		$f(X)$	N	v	R^2	Residual distribution plot	P_{rc} ($\pm$ 1%)	P_{rc} ($\pm$ 5%)	P_{rc} ($\pm$ 10%)	SD	AAE	ARE[a]
1	Normal boiling point [K]	$exp\left(\frac{T_b}{T_{bo}}\right)$	3510	3496	0.8865		23.85	76.10	95.75	25.54	19.51	4.48
2	Critical temperature [K]	$exp\left(\frac{T_c}{T_{co}}\right)$	858	844	0.8592		27.74	72.26	93.47	39.82	29.24	4.70
3	Critical pressure [bar]	$\left(P_c-P_{c1}\right)^{-0.5}-P_{c2}$	852	838	0.8342		7.86	47.77	84.15	6.94	4.74	11.35
4	Critical volume [cc/mol]	V_c-V_{c0}	797	783	0.9908		34.63	88.96	97.87	18.48	12.16	3.12

5	Normal melting point [K]	$exp\left(\frac{T_m}{T_{mo}}\right)$	5183	5169	0.7135		8.47	41.75	71.77	51.31	38.68	12.32
6	Gibbs free energy [kJ/mol]	$\Delta_f G^o - G_{f0}$	749	737	0.9555		2.94	18.69	41.52	44.90	32.23	**
7	Enthalpy of formation [kJ/mol]	$\Delta_f H^o{}_{gas} - H_{f0}$	882	869	0.9271		4.65	27.21	49.32	67.39	44.70	**
8	Enthalpy of fusion [kJ/mol]	$\Delta_{fus} H - H_{fus0}$	764	751	0.6913		4.32	18.72	35.34	7.00	4.04	**
9	Octanol/Water partition coefficient	$LogK_{ow} - K_{ow0}$	12193	12179	0.6705		2.61	14.68	28.76	1.03	0.78	**

10	Flash point [K]	$F_p - F_{po}$	512	500	0.7951		31.45	86.91	97.46	14.30	10.66	3.27
11	Enthalpy of vaporization (298K) [kJ/mol]	$\Delta_{vap}H° - H_{v0}$	705	692	0.7232		8.51	43.12	66.52	7.31	4.63	10.47
12	Enthalpy of vaporization (Tb) [kJ/mol]	$\Delta_{vap}H - H_{vb0}$	512	499	0.5668		17.19	51.56	77.73	4.72	2.98	8.14
13	Acentric factor	$\exp\left(\frac{\omega}{\omega_a}\right)^{\omega_b} - \omega_c$	1723	1709	0.6243		4.24	23.33	40.57	0.1837	0.1207	25.74
14	Liquid molar volume [cc/kmol]	$V_m - V_{m0}$	1056	1042	0.9918		35.13	85.51	94.98	0.007	0.004	3.66

[a] For $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, and $LogK_{ow}$, ARE is not reported since these properties contain both positive and negative values.

The variables T_{b0}, T_{c0}, P_{c1}, P_{c2}, V_{c0}, T_{m0}, G_{f0}, H_{f0}, H_{fus0}, K_{ow0}, F_{p0}, Ait_1, Ait_2, H_{v0}, H_{vbo}, S_{vb0}, δ_0, ω_a, ω_b, ω_c, and V_{mo} as defined in the functional forms, $f(X)$ given in Tables 4.2 and 4.3 are additional adjustable parameters of property prediction models. The values of these parameters together with the uncertainty in terms of 95% confidence interval are listed in Table 4.4. The large differences in the value of V_{c0}, G_{f0}, and H_{f0} in step-wise and simultaneous method are due to different approach employed for the parameter regression. In step-wise regression method, the universal parameter is regressed in the first level estimation involving only first-order groups while in simultaneous regression method, the universal parameter is regressed together with first, second, and third-order groups. The total list of groups and their contributions C_i, D_j, and E_k for the 21 pure component properties modeled in this work are given in the Appendix A (see Tables A2-A4 for MG method based models analysed using step-wise regression method, and Tables A5-A7 for MG method based models analysed using simultaneous regression method). The list of atoms, their contributions a_i, adjustable model parameters (b and c), and universal parameter d for CI method based property prediction models are given in the Appendix A (see Table A8).

Table 4.4 Values of universal constants (additional adjustable parameters) [a]

Universal constants	Step-wise method		Simultaneous method	
	Value	Uncertainty (95% CI)	Value	Uncertainty (95% CI)
T_{b0} [K]	244.51	0.30	244.78	0.28
T_{c0} [K]	181.67	0.60	181.67	0.72
P_{c1} [bar]	0.0519	0.002	0.0519	0.001
P_{c2} [$bar^{-0.5}$]	0.1347	0.004	0.1100	0.003
V_{c0} [cc/mol]	28.00	1.45	14.61	3.06
T_{m0} [K]	143.57	1.6	144.09	1.81
G_{f0} [kJ/mol]	-1.33	1.19	8.50	2.95
H_{f0} [kJ/mol]	35.17	13.3	83.96	29.76
H_{v0} [kJ/mol]	9.61	1.4	10.43	4.43
H_{fus0} [kJ/mol]	-1.77	0.37	-1.29	0.12
K_{ow0} [-]	0.48	0.04	0.75	0.05
F_{p0} [K]	170.7	11.8	150.0	31.19
Ait_1 [b] [-]	-	-	71.25	10.28
Ait_2 [b] [K]	-	-	525.93	134.1
H_{vbo} [kJ/mol]	15.4	1.25	15.1	3.1
S_{vb0} [kJ/mol]	83.30	2.58	83.77	6.7
δ_0 [$MPa^{1/2}$]	21.66	0.78	20.73	1.3
ω_a [-]	0.90	0.004	0.91	0.002
ω_b [-]	0.10	0.003	0.04	0.0001
ω_c [-]	1.001	0.12	1.003	0.17
V_{mo} [cc/kmol]	0.016	0.002	0.012	0.003

[a] The values of universal constants for the CI models are the same as those given for the step-wise method.

[b] Ait property is modeled using simultaneous regression method only.

4.1.3 Selection of quantity of property data-set for parameter regression

It is important to note that all of the available experimental data of the pure component properties have been considered for the parameter regression purpose. By including all of the available experimental data of the property in the regression, it is possible to improve the prediction accuracy and application range of GC models. To illustrate this point, for standard enthalpy of formation, the effect of quantity of experimental data on the quality of the parameter estimation is shown by considering different combinations of training sets and validation sets formed by random selection of the experimental data-points (see Table 4.5). It can be seen that better model performance statistics (lower SD, and lower AAE) is obtained for the case in which all of the experimental data-points are considered for the regression purpose. Also, the total number of model parameters involved in the regression (i.e. model parameters for which estimation of contribution was possible) is highest when all of the data-points are considered in the regression.

Table 4.5 Effect of quantity of experimental data on the quality of parameter estimation

Distribution of experimental data	SD in kJ/mol	AAE in kJ/mol	No. of model parameters estimated [a]
50% for training purpose	9.59	5.31	194
66.67% for training purpose	8.73	4.75	217
75% for training purpose	7.09	4.34	229
All data-points for training purpose	6.60	4.15	232

[a] Total no. of model parameters (first-order, second-order and third-order MG groups) is 424.

4.1.4 Reliability of developed property models

The reliability of the developed MG method based property models has been tested by comparing model prediction uncertainties with reported range of experimental measurement uncertainties for the properties with related available data (see Table 4.6). The DIPPR 801® databank has been the main source for experimental measurement uncertainties for the listed properties in Table 4.6. The DIPPR 801® database provides available data source together with the quality codes containing measurement uncertainty in the property values in % terms (for e.g. <1%, <3%, <10% etc). To get an average measurement uncertainty value, the limiting values (<1% as 1%, <3% as 3% etc.) are taken, to calculate the average measurement uncertainties based on the reported experimental values of the properties. The data-points given in Table 4.6 indicates the number of experimental data-points for which quality code

(and hence measurement uncertainty) was available in the DIPPR 801® database. It is to be noted that, the upper limit of the uncertainty in the measurement is taken into consideration and therefore, the actual average measurement uncertainty will be lower than the calculated measurement uncertainties shown in Table 4.6. However, this comparison helps to provide an indication of the quality (and hence reliability) of the predictions and it can be noticed that for most of the properties the prediction error is lower than (or at least comparable to) the average measurement error, except for the case of normal melting point and standard enthalpy of fusion. For these two properties, group contribution methods, in general, have difficulties in providing a reliable estimation. This is mainly due to the strong dependency of melting point on intermolecular interaction and molecular symmetry (Poling et al., 2000).

Table 4.6 Comparison of model prediction error with reported average measurement error

Property	Data-points	Average measurement error [a]	Average prediction error [b]
Normal boiling point [K]	1306	6.32	6.17
Critical temperature [K]	402	7.95	7.72
Critical pressure [bar]	293	1.20	1.40
Critical volume [cc/mol]	234	22.30	7.97
Normal melting Point [K]	1385	5.10	15.99
Gibbs free energy [kJ/mol]	258	4.60	5.24
Enthalpy of formation [kJ/mol]	668	6.31	5.03
Enthalpy of fusion [kJ/mol]	520	0.47	2.79
Flash point [K]	111	27.96	8.97

[a] The listed measurement errors are taken from the DIPPR 801® databank

[b] The listed prediction errors are from step-wise regression method

4.1.5 Thermodynamic consistency and predictive power

The thermodynamic consistency of predicted property values and the predictive power of the developed GC models are tested by studying the behavior of certain properties of *n*-alkanes as the carbon number goes to a very high number. The following discussion is based on the properties predicted using MG method based property models analysed using step-wise method for the parameter regression.

4.1.5.1 Relation of T_b versus T_c for n-alkanes

A plot of predicted values of T_b and T_c of *n*-alkanes versus the increasing carbon number of *n*-alkanes is shown in Figure 4.1. The experimental data available for T_b and T_c for *n*-alkanes are also shown in the same plot. In agreement with the basic physical principles, throughout

the homologous series, the ratio of T_c/T_b is greater than unity (Tsonopoulos, 1987; Tsonopoulos and Tan, 1993; Constantinou and Gani, 1994).

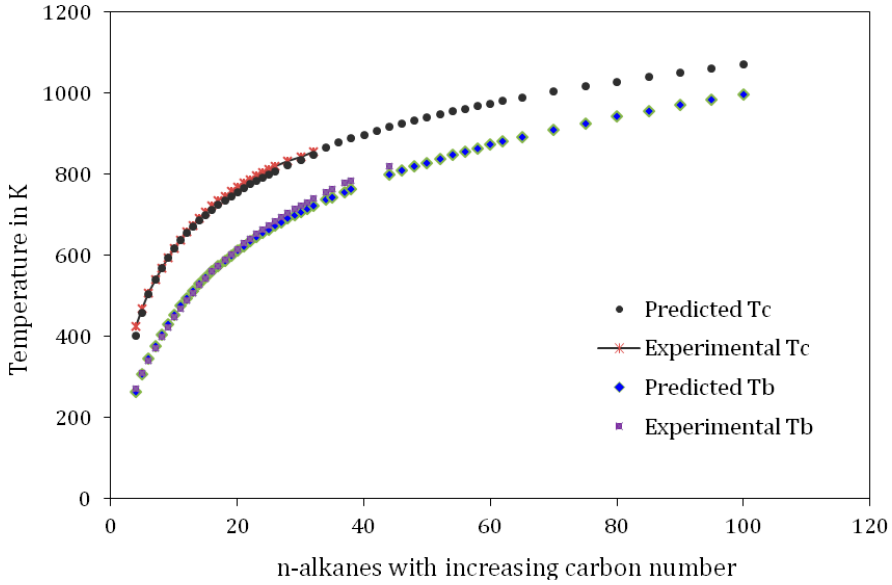


Figure 4.1 T_b versus T_c for *n*-alkanes

The relation of T_b versus T_c was also tested for a data-set (850 components) comprising a wide range of organic components (hydrocarbons, oxygenated, nitrogenated halogenated and multifunctional types) and it was observed that the ratio T_c/T_b is greater than 1 for each pure component present in the data-set (see Figure 4.2(a)).

4.1.5.2 The ratio of T_c / P_c of n-alkanes

The ratio T_c/P_c is important in many engineering calculations involving use of the equation-of-state based property models and hence reliable predictions of this ratio is important. Figure 4.2(b) shows a plot of predicted values of T_c/P_c versus experimental values of T_c/P_c (up to carbon number 32) and it can be seen that this ratio is predicted with a high degree of accuracy and reliability.

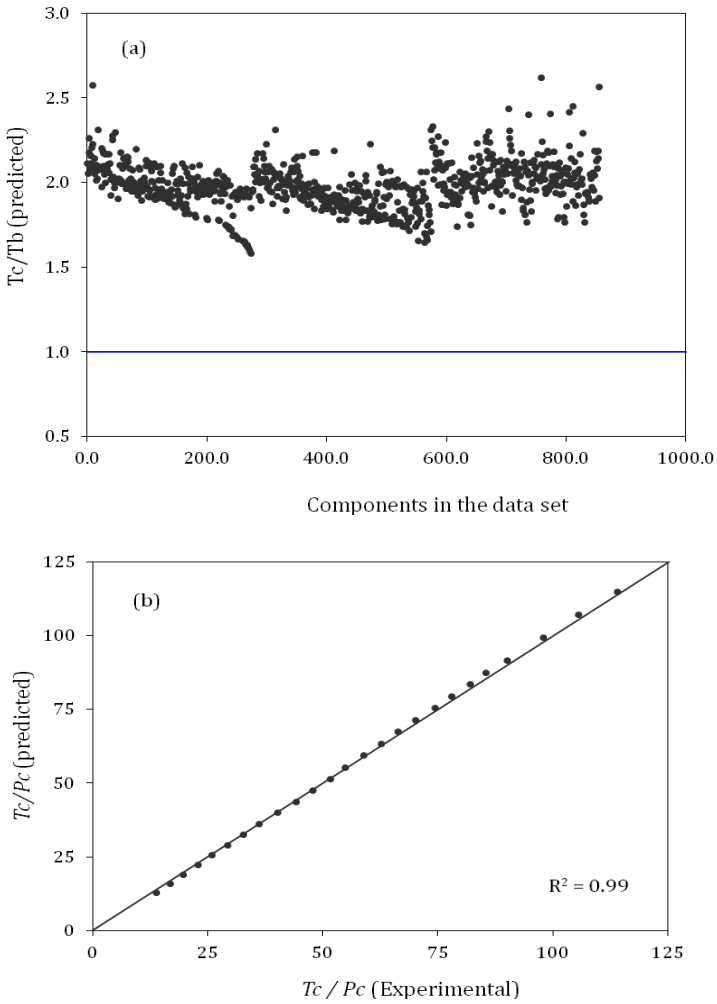


Figure 4.2 Plots for: (a) T_c/T_b ratio for a wide range of components in the data set; (b) Experimental values of T_c/P_c versus predicted values of $T_c/P_{c.}$

4.1.5.3 Critical pressure of n-alkanes

Tsonopoulos and Tan (1993) have reported two different lower limits for critical pressure which are: $P_{c\infty}$ = 2.68 bar and $P_{c\infty}$ = 0.0519 bar based on two different experimental data-sets of critical constants. The limit $P_{c\infty}$ = 0.0519 bar is used in regressing the data of critical pressure. Figure 4.2(c) shows a plot of predicted values of P_c of *n*-alkanes and it can be seen that as carbon number goes to a high number, the critical pressure approaches a minimum which is consistent with the reasoning that a very large molecule (with infinite chain length) cannot exists as a vapor (Tsonopoulos, 1987).

4.1.5.4 Critical compressibility factor (Z_c) of n-alkanes

McFarlane et al. (2010) employed MG method based property model of T_c, P_c and V_c to calculate Z_c for *n*-alkanes with molecular weight up to 4500 g/mol and found that the calculated Z_c exceeds the limiting value as given by Poling et al. (2000) (Z_c should be less than 0.291 for molecules with critical temperature greater than 100 K) for molecular weight above about 1200 g/mol. The higher values of Z_c (> 0.291) was mainly due to the higher limiting value of P_c (5.99 bar) predicted by the MG method (2001). For this reason, the limiting value of $P_{c\infty}$ = 0.0519 bar was used in regressing the data of critical pressure to obtain more reliable estimation of critical pressure of *n*-alkanes with large molecular weight and thus obtain Z_c values in consistent with the theoretical foundation. The calculated Z_c values using RT_c/P_cV_c (with R =83.14 cc-bar/mol-K, T_c in K, P_c in bar and V_c in cc/mol) for *n*-alkanes with the revised and improved model parameters of T_c, P_c, and V_c are shown in Figure 4.2(d).

4.1.5.5 Critical density (d_c) of n-alkanes

Critical density can be expressed as an inverse of the critical volume. An *n*-alkane molecule with c number of carbons can be represented by CH_3 group with 2 occurrences and by CH_2 group with c-2 occurrences. The molecular weight of *n*-alkane can be represented by $14c+2$ (kg/kmol). Therefore, expression for d_c (in kg/m^3) after substituting values of group-contributions for CH_3 and CH_2 is given by:

$$d_c = \frac{1}{V_c} = \frac{(14c+2)}{(56.5948c-24.0164)} \times 1000 \tag{4.1}$$

As c tends to an infinite value, the limiting value for the critical density is 247.37 kg/m^3. This calculated limiting value for critical density seems to be reasonable and in agreement with the value discussed by Tsonopoulos and Tan (1993).

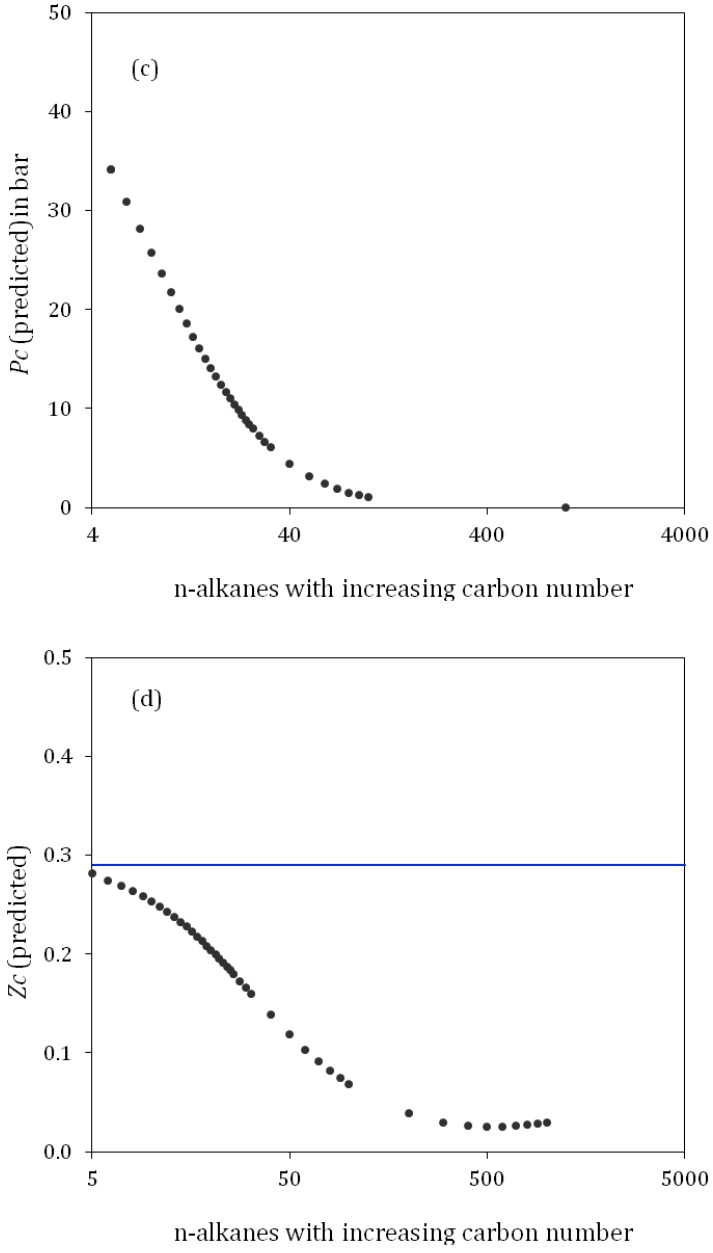
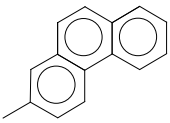


Figure 4.2 Plots for: (c) Predicted P_c of *n*-alkanes; (d) Predicted Z_c of *n*-alkanes.

4.1.6 Application example

The application of the developed models to estimate properties of pure components and quantify uncertainties of estimated property values is illustrated by considering predictions of T_b using the model parameters obtained from simultaneous regression method for Phenanthrene, 2-methyl- (CAS No. 2531-84-2). Table 4.7 provides information of first-order, second-order and third-order MG groups used to represent Phenanthrene, 2-methyl-, their frequency and contributions for each group (T_{b1i}, T_{b2j}, and T_{b3k}) taken from Appendix A (see Tables A5-A7). Using this information and the property model for T_b, we estimate the T_b of Phenanthrene, 2-methyl- as 619.1 K (Experimental value is 623.1 K).

Table 4.7 Estimation of T_b of Phenanthrene, 2-methyl-

Phenanthrene, 2-methyl-	**Molecular structure**	
CAS No. 2531-84-2		
Molecular formula: $C_{15}H_{12}$		
First-order groups	**Occurrences**	**Contribution**
aCH	9	0.7332
aC fused with aromatic ring	4	1.2531
aC-CH_3	1	1.2616
Second-order groups	**Occurrences**	**Contribution**
	No second-order groups are involved	
Third-order groups	**Occurrences**	**Contribution**
AROM.FUSED[2]	1	-0.1599
AROM.FUSED[2]s^2	1	-0.1829
AROM.FUSED[4p]	1	0.0119

$$T_b^{pred} = T_{bo}\, ln\left(\sum_i N_i C_i + w\sum_j M_j D_j + z\sum_k E_k O_k\right) = 619.1 \text{ K}$$ Absolute error = 623.1 – 619.1 = 4.0 K

As discussed in chapter 3, to quantify the uncertainties of estimated T_b, we use information of covariance $COV(\boldsymbol{P}^*)$ of the involved groups and universal constant T_{bo}, and local sensitivity $J(\boldsymbol{P}^*)$ of the T_b model. The covariance of the listed groups as given in Table 4.8 was extracted from the whole covariance matrix for all the groups obtained for the case of T_b model analysed using simultaneous regression method. In Table 4.8, only lower triangular elements are shown since the upper triangular matrix elements are identical to the lower one. Table 4.9 lists the local sensitivity of the T_b model with respect to the model parameters (for contributions listed in Table 4.7 and universal constant T_{bo}).

Table 4.8. Covariance matrix $COV(\boldsymbol{P}^*)$ with dimensions (7×7) for groups listed in Table 4.7

	T_{bo}	aCH	aC	aC-CH_3	AROM.FUSED[2]	AROM.FUSED[2]s^2	AROM.FUSED[4p]
T_{bo}	0.0781						
aCH	-0.00028	2.9E-05					
aC	-0.0017	7.0E-06	0.0101				
aC-CH_3	-0.00079	-1.4E-05	-8.1E-05	0.00021			
AROM.FUSED[2]	3.79E-05	-0.0001	-0.0112	0.00020	0.01465		
AROM.FUSED[2]s^2	0.00017	-5.3E-05	-0.0101	4.58E-06	0.011201	0.01262	
AROM.FUSED[4p]	-8.84E-05	-7.0E-05	-0.0192	0.00019	0.020194	0.01936	0.06496

Table 4.9. Local sensitivity $J(\boldsymbol{P}^*)$ with dimensions (1×7) of T_b model with respect to model parameters

$\delta T_b/\delta T_{bo}$	$\delta T_b/\delta$aCH	$\delta T_b/\delta$aC	$\delta T_b/\delta$aC-CH_3	$\delta T_b/\delta$ AROM.FUSED[2]	$\delta T_b/\delta$ AROM.FUSED[2] s^2	$\delta T_b/\delta$ AROM.FUSED[4p]
2.5290	175.6577	78.07	19.5175	19.5175	19.5175	19.5175

To calculate the confidence intervals, say the 95% confidence intervals of the estimated T_b value, the covariance matrix $COV(\boldsymbol{P}^*)$ given in Table 4.8 and the local sensitivity $J(\boldsymbol{P}^*)$ given in Table 4.9 are substituted in Eq. (3.13). For 95% confidence interval calculation, the t-distribution value corresponding to 0.05 /2 percentile (i.e. α_t /2 percentile) and with 3179 degrees of freedom (taken from Table 4.2) is obtained by solving Eq. (3.12) for t and this value is 1.9607. The predicted value of the normal boiling point $T_b^{\,pred}$ is 619.1 K (see Table 4.7). The calculated 95% confidence intervals of the estimated $T_b^{\,pred}$ value is therefore,

$$T_{b\,(1-0.05)}^{\,pred} = \underbrace{T_b^{\,pred}}_{619.1} \pm \underbrace{\sqrt{\mathrm{diag}\left(J\left(\boldsymbol{P}^*\right)COV\left(\boldsymbol{P}^*\right)J\left(\boldsymbol{P}^*\right)^T\right)}}_{3.30} \cdot \underbrace{t\left(\nu,\alpha_t/2\right)}_{1.9607} = 619.1\ \mathrm{K} \pm 6.5\ \mathrm{K}$$

It can be observed that the experimental value of T_b (623.1 K) lies within the predicted confidence intervals indicating reliability of the developed methodology for estimating the property values and uncertainties in the estimated property values. This, of course, can only be checked when experimental data is available. However, if uncertainty estimate of the measured data is available, then if the calculated 95% confidence interval is smaller, the estimated property value could be considered as safe. This is further illustrated in Figure 4.3 by plotting experimental values of T_b and the calculated confidence intervals at 95% confidence level (shown as vertical bars) for the components in the dataset and it can be seen from Figure 4.3 that the most of the experimental values falls within the calculated confidence intervals. This analysis supports both the linear error propagation method used for quantifying the model prediction error and quality of the resulting estimated property values and their associated confidence intervals.

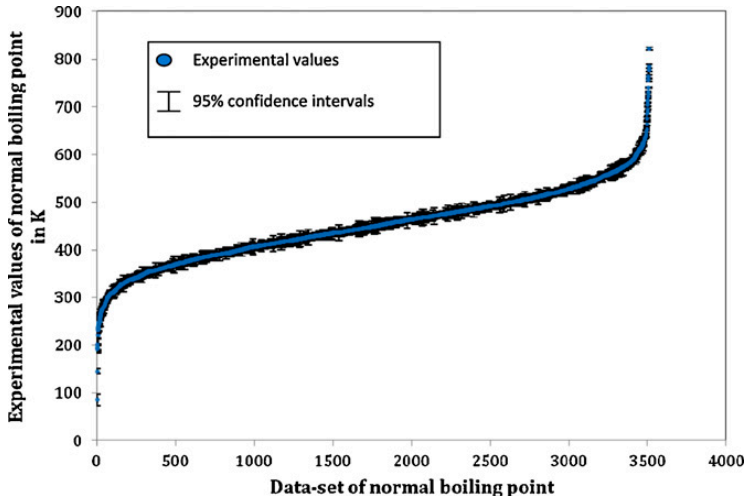


Figure 4.3 Experimental values of T_b and the calculated 95% confidence intervals for the pure components in the data-set

Note that in order to simplify the illustration of the application of the methodology, the calculation of confidence intervals of the estimated property values using models analysed by simultaneous regression method is shown here since there will be a single covariance matrix containing covariance of all the listed groups and parameters. The approach discussed in this section to quantify the uncertainties in the property value is the same for the case of property models analysed using the step-wise regression method. In the case of step-wise regression method, there will be a covariance matrix for each order of the groups, i.e., first-order groups, second-order groups and third-order MG groups and hence, quantification of uncertainty in

the predicted property value is to be performed (using these covariance matrices) after each step (level-1, level-2 and level-3 as discussed in the section 3.2 of chapter 3.

4.1.7 Summary

Property modeling and analysis of GC^+ based models for thermo-physical and transport-related pure component properties have been performed to provide more reliable predictions of these properties together with an estimate of prediction errors (uncertainties). To improve the performance, the application range, and the predictive power of earlier version of the property models, an updated CAPEC database containing large experimental data-sets of properties is used in the parameter estimation to develop revised and improved parameters of GC^+ models. The application range of GC^+ models is increased by making use of the new experimental data available in the CAPEC database. In addition, simultaneous regression approach for the estimation of model parameters of MG method based property models is developed and used. In total, 21 pure component properties have been modeled and analysed. The acentric factor and the liquid molar volume (298) which were modeled earlier by the Constantinou and Gani method (1995) are modeled using the MG method to provide improved accuracy and the application range of the models. The application of the developed GC^+ models to estimate pure component properties together with uncertainties of predicted property values is highlighted through an application example. The reliability of developed property models has been tested and illustrated by comparing model prediction uncertainties with reported range of experimental measurement uncertainties for the properties with related available data. Important issues related to property modeling such as thermodynamic consistency of the predicted properties, and predictive power of the models has been addressed. The developed models for the estimation of thermo-physical and transport-related properties are simple, yet sound and effective and provides not only the estimated property values but also the uncertainties in the estimated property values. This feature would allow one to evaluate effects of these uncertainties on the product-process design involving use of predictive models thereby contributing to better-informed and reliable engineering solutions.

4.2 Development of GC^+ Models for Environmental-Related Properties

Motivated by the need of reliable estimation of environmental-related properties in synthesis, design, and analysis of sustainable processes, this work aims to develop GC^+ approach based property models to provide reliable estimation of environmental-related properties together with uncertainties of estimated property values. For this purpose, the methods for performing property modeling as discussed in chapter 3 are used. In total 22 environmental-related properties, which include the fathead minnow 96-h LC_{50} (LC_{50}(FM)), *daphnia magna* 48-h LC50 (LC_{50}(DM)), oral rat LD_{50}, aqueous solubility ($LogW_s$), bio-concentration factor (BCF), permissible exposure limit (PEL(OSHA-TWA)), photochemical oxidation potential (PCO), global warming potential (GWP), ozone depletion potential (ODP), and acidification potential (AP), emission to urban air (carcinogenic (EUA_C) and non-carcinogenic (EUA_{NC})), emission to continental rural air (carcinogenic (ERA_C) and non-carcinogenic (ERA_{NC})), emission to continental fresh water (carcinogenic (EFW_C) and non-carcinogenic (EFW_{NC})), emission to continental sea water (carcinogenic (ESW_C) and non-carcinogenic (ESW_{NC})), emission to continental natural soil (carcinogenic (ENS_C) and non-carcinogenic (ENS_{NC})), emission to continental agricultural soil (carcinogenic (EAS_C) and non-carcinogenic (EAS_{NC})) have been modeled and analysed. The definition of these properties is given in Table 4.10.

Table 4.10 Definition of environmental-related properties considered in this work

Sl. No.	Property	Definition of the property
1	Fathead Minnow 96- hr LC_{50} (LC_{50} (FM)) in mol/lit	The fathead minnow LC_{50} endpoint represents the concentration in water which kills half of fathead minnow (*Pimephales promelas*) in 4 days (96 hours) (US EPA, T.E.S.T., 2012).
2	*Daphnia magna* 48-hr LC_{50} (LC_{50} (DM)) in mol/lit	The *Daphnia magna* LC_{50} endpoint represents the concentration in water which kills half of *Daphnia magna* (a water flea) in 48 hours (US EPA, T.E.S.T., 2012).
3	Oral rat LD_{50} (LD_{50}) in mol/kg	The oral rat LD50 endpoint represents the amount of the chemical (mass of the chemical per body weight of the rat) which when orally ingested kills half of rats (US EPA, T.E.S.T., 2012).
4	Aqueous solubility ($LogW_S$) in gm/lit	Aqueous solubility is defined as the amount of a chemical that will dissolve in liquid water to form a homogeneous solution (US EPA, T.E.S.T., 2012).
5	Bio-concentration factor (BCF)	The bio-concentration factor is defined as the ratio of the chemical concentration in biota as a result of absorption via the respiratory surface to that in water at steady state (US EPA, T.E.S.T., 2012).
6	Permissible exposure limit (OSHA-TWA) in mol/m3	The permissible exposure limit (OSHA-TWA) is a legal limit in the United States for exposure of an employee to a chemical substance or physical agent. It is usually given as a time-weighted average (TWA). A TWA is the average exposure over a specified period of time, usually a nominal eight hours. This means that, for limited periods, a worker may be exposed to concentrations higher than the permissible exposure limit, so long as the average concentration over eight hours remains lower (OSHA PEL http://www.osha.gov/SLTC/pel/).
7	Photochemical oxidation potential (PCO)	The photochemical oxidation potential is the result of reactions that take place between nitrogen oxides and volatile organic components exposed to UV radiation. It is expressed using a reference substance such as ethylene (Heijungs et al., 1992).

8	Global warming potential (GWP)	The global warming potential is calculated as a sum of emissions of the greenhouse gases (CO_2, N_2O, CH_4 and VOCs) multiplied by their respective GWP factors (Heijungs et al., 1992).
9	Ozone depletion potential (ODP)	The ozone depletion potential is defined as the ozone depletion produced by a unit of the gas converted into ozone depletion values produced by the reference substance trichlorofluoromethane (Heijungs et al., 1992).
10	Acidification potential (ODP)	The acidification potential is a measure of the disposition of a unit of the mass of a component to release H^+ protons, expressed in terms of the H^+ potential of the reference substance SO_2 (Heijungs et al., 1992).
11	Emission to urban air (EUA_C) in cases/kg emitted (carcinogenic)	Estimated increase in morbidity (carcinogenic) in the total human population per unit mass of a chemical emitted in the urban air compartment (that is, emission to higher population density). (USEtox™ model. http://www.usetox.org/)
12	Emission to urban air (EUA_{NC}) in cases/kg emitted (non-carcinogenic)	Estimated increase in morbidity (non-carcinogenic) in the total human population per unit mass of a chemical emitted in the urban air compartment (that is, emission to higher population density). (USEtox™ model. http://www.usetox.org/)
13	Emission to continental rural air (ERA_C) in cases/kg emitted (carcinogenic)	Estimated increase in morbidity (carcinogenic) in the total human population per unit mass of a chemical emitted in the rural air compartment (that is, emission to lower population density, lower stratosphere, and upper troposphere). (USEtox™ model. http://www.usetox.org/)
14	Emission to continental rural air (ERA_{NC}) in cases/kg emitted (non-carcinogenic)	Estimated increase in morbidity (non-carcinogenic) in the total human population per unit mass of a chemical emitted in the rural air compartment (that is, emission to lower population density, lower stratosphere, and upper troposphere). (USEtox™ model. http://www.usetox.org/)
15	Emission to continental fresh water (EFW_C) in cases/kg emitted(carcinogenic)	Estimated increase in morbidity (carcinogenic) in the total human population per unit mass of a chemical emitted in the fresh water compartment (that is, emission to lakes, river, and ground-water). (USEtox™ model. http://www.usetox.org/)

16	Emission to continental fresh water (EFW_{NC}) in cases/kg emitted (non-carcinogenic)	Estimated increase in morbidity (non-carcinogenic) in the total human population per unit mass of a chemical emitted in the fresh water compartment (that is, emission to lakes, river, and ground-water). (USEtox™ model. http://www.usetox.org/)
17	Emission to continental sea water (ESW_C) in cases/kg emitted (carcinogenic)	Estimated increase in morbidity (carcinogenic) in the total human population per unit mass of a chemical emitted in the sea water compartment. (USEtox™ model. http://www.usetox.org/)
18	Emission to continental sea water (ESW_{NC}) in cases/kg emitted (non-carcinogenic)	Estimated increase in morbidity (non-carcinogenic) in the total human population per unit mass of a chemical emitted in the sea water compartment. (USEtox™ model. http://www.usetox.org/)
19	Emission to continental natural soil (ENS_C) in cases/kg emitted (carcinogenic)	Estimated increase in morbidity (carcinogenic) in the total human population per unit mass of a chemical emitted in the natural soil compartment (that is, emission to forestry, and industrial soil). (USEtox™ model. http://www.usetox.org/)
20	Emission to continental natural soil (ENS_{NC}) in cases/kg emitted (non-carcinogenic)	Estimated increase in morbidity (non-carcinogenic) in the total human population per unit mass of a chemical emitted in the natural soil compartment (that is, emission to forestry, and industrial soil). (USEtox™ model. http://www.usetox.org/)
21	Emission to continental agricultural soil (EAS_C) in cases/kg emitted (carcinogenic)	Estimated increase in morbidity (carcinogenic) in the total human population per unit mass of a chemical emitted in the agricultural soil compartment. (USEtox™ model. http://www.usetox.org/)
22	Emission to continental agricultural soil (EAS_{NC}) in cases/kg emitted (non-carcinogenic)	Estimated increase in morbidity (non-carcinogenic) in the total human population per unit mass of a chemical emitted in the agricultural soil compartment. (USEtox™ model. http://www.usetox.org/)

4.2.1 Database

For the estimation of property model parameters, experimental data-sets of organic chemicals of various classes from the database of US Environmental Protection Agency (EPA) and from the database of USEtox^{TM} are used. The details of data-set of each property are given in Table 4.11(a) for the US EPA database and in Table 4.11(b) for the USEtox^{TM} database.

Table 4.11(a) Description of the US EPA data-sets used for the regression purpose

class of chemicals	LC_{50} (FM)	LC_{50} (DM)	LD_{50}	$LogW_s$	BCF	PCO	PEL	GWP	ODP	AP
Hydrocarbons	32	19	69	236	79	337	98	0	0	0
Oxygenated	238	54	1382	1110	76	244	127	1	0	0
Nitrogenated	80	24	397	244	57	8	45	0	0	0
Chlorinated	48	37	111	274	77	23	41	5	3	5
Fluorinated	1	0	3	21	1	5	4	23	0	0
Brominated	10	4	14	47	15	5	7	2	1	0
Iodinated	1	0	5	17	0	0	1	0	0	0
phosphorous containing	0	0	5	0	0	0	0	0	0	0
Sulfonated	9	8	24	19	5	0	15	0	0	0
silicon containing	0	0	1	2	0	0	0	0	0	0
multifunctional	390	174	3984	2711	352	17	87	20	24	5
total number of chemicals	809	320	5995	4681	662	639	425	51	28	10

Table 4.11(b) Description of the USEtox^{TM} data-sets used for the regression purpose

class of chemicals	EUA_C	EUA_{NC}	ERA_C	ERA_{NC}	EFW_C	EFW_{NC}	ESW_C	ESW_{NC}	ENS_C	ENS_{NC}	EAS_C	EAS_{NC}
Hydrocarbons	25	14	18	16	19	14	19	16	18	16	20	16
Oxygenated	107	56	96	60	98	57	101	60	96	58	97	58
Nitrogenated	29	14	27	14	27	13	26	15	27	15	27	14
Chlorinated	46	23	43	26	44	27	45	32	45	30	43	28
Fluorinated	4	1	4	1	4	1	4	1	4	1	4	1
Brominated	6	2	5	2	4	2	5	2	5	3	5	3
Iodinated	0	0	0	0	0	0	0	0	0	0	0	0
phosphorous	0	0	0	0	0	0	0	0	0	0	0	0
Sulfonated	3	1	3	1	3	1	3	1	3	1	3	1
silicon containing	0	0	0	0	0	0	0	0	0	0	0	0
multifunctional	236	230	274	229	273	230	274	233	262	238	271	231
total number of chemicals	456	341	470	349	472	345	477	360	460	362	470	352

In following sections, the selection of suitable property models for modeling environmental-related properties and the performance statistics for the developed property models are discussed. The results are presented for the following pure component property prediction models:

- MG method based property models analyzed using step-wise regression method
- MG method based property models analyzed using simultaneous regression method
- CI method based property models

4.2.2 Selection of suitable property models for environmental-related properties

In this work, the basis for selecting an appropriate property model for the environmental-related property has been the study of behavior of that property of certain class of chemicals with increasing molecular weight. This is illustrated for the case of LC_{50} (FM). Figure 4.4 shows plots of various molecules types with increasing molecular weight versus their experimental values of −Log LC_{50} (FM). It can be seen that this plot is almost a straight line suggesting that the property LC_{50} (FM) can be modeled using a linear model of the form,

$$-\text{Log}LC_{50}(\text{FM}) + \text{Constant} = \sum_i N_i C_i + w\sum_j M_j D_j + z\sum_k E_k O_k \qquad (4.2)$$

Similar analysis has been performed to obtain a suitable form of the property model for other environmental-related properties to decide the form of the property model function so that accurate and reliable estimation of environmental-related properties can be achieved.

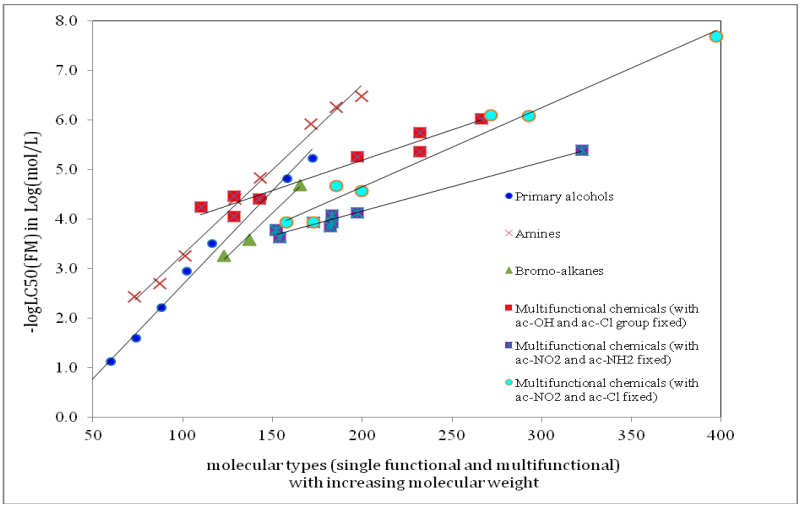


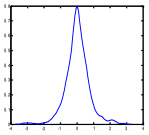
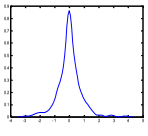
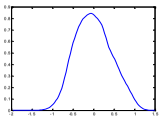
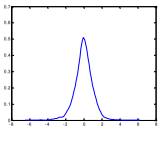
Figure 4.4 Plot of molecule types versus their experimental values of –Log LC_{50}(FM)

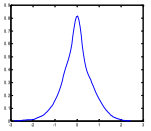
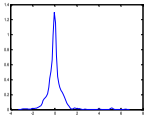
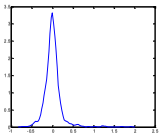
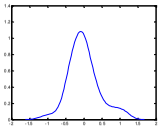
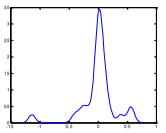
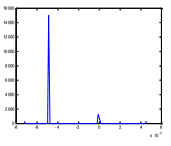
4.2.3 Model performance statistics

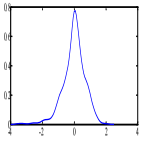
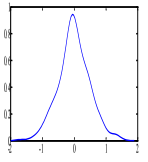
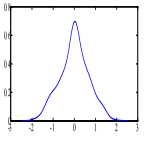
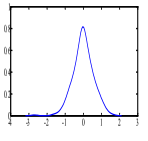
The model performance statistics for property models analysed using the step-wise regression method are provided in Table 4.12. The model performance statistics for properties analysed using the simultaneous regression method are given in Appendix B (see Table B1). The model performance statistics for property models analysed using the CI method are provided in Table 4.13. The property models developed based on the CI method (see Table 4.13) have reasonable model performance statistics. Marrero and Gani (2002) reported SD, AAE, and R^2 values for the GC model for $LogW_s$ as 0.55, 0.46 and 0.93 respectively. In their analysis, the number of estimated model parameters (groups) are 155 first order groups, 99 second order groups, and 48 third-order groups (that is, total 302 groups estimated out of 424 groups). Referring to the Table 4.12, it can be seen that the property model for $LogW_s$ has SD, AAE, and R^2 values of 0.99, 0.73, and 0.78 respectively. In this work, the number of estimated groups is 197 first order groups, 124 second order groups, and 57 third-order groups (total 378 groups estimated out of 424 groups). It is to be noted that in the present work, a much larger data-set (4681 data-points as compared to 2087 data-points used by Marrero and Gani, 2002) of $LogW_s$ comprising complex and polyfunctional environmental-related chemicals is used in the regression, which makes it possible to estimate larger number of model parameters thereby contributing to improved application range of the property model for $LogW_s$. A similar note can be made for the case of property model for LC_{50}(FM). The developed property model for LC_{50}(FM) has SD, AAE, and R^2 values of 0.69, 0.48, and 0.78 respectively. Martin and Young (2000) reported SD and R^2 values for their GC model for LC_{50}(FM) as 0.37 and 0.91 respectively. The use of the large data-set for LC_{50}(FM) allows estimation of a large number of model parameters which in turn allows one to estimate LC_{50} (FM) for a wide range of organic chemicals. For the property LC_{50}(DM), the model performance statistics are similar to that of LC_{50}(FM) model. The developed property model for LD_{50} using a data-set of 5995 chemicals has reasonably good performance statistics with SD, AAE, and R^2 values as 0.43, 0.35 and 0.73 respectively. Several estimation methods based on the QSAR approach have been reported in the literature that uses other properties such as LC_{50}(DM) as an input to their estimation method to estimate LD_{50}. Also, these methods have been developed based on relatively smaller data-sets (with few hundreds of chemicals in the data-set). The application of such methods is restricted by the availability of the experimental data of the needed input properties for their estimation. A similar issue is associated with the estimation methods for BCF requiring additional inputs such as the octanol/water partition coefficient. In this work, the developed property model for BCF has SD, AAE, and R^2 values of 0.63, 0.47, and 0.78 respectively. It is to be noted that the developed property models for LD_{50} and for BCF only require the molecular structure of the chemical for the property estimation. From Table 4.12 and 4.13, it can be seen that for some properties such as, photochemical oxidation, emission to urban air (carcinogenic and

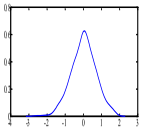
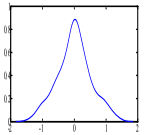
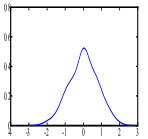
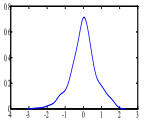
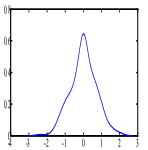
noncarcinogenic) etc, the residual distribution do not follow normal distribution. This non-guassian distribution could be due to the nature of the property model. For properties GWP, ODP, and AP the number of experimental data points used in the regression are smaller as compared to other properties analysed in this work. However, it can be noted that these properties belong to a particular class of chemicals. For example, global warming potential and ozone depletion potential properties involve halogenated chemicals, and acidification potential property involves nitrogenated chemicals. Hence, even though the experimental data-sets for these properties are smaller, the models for these properties are able to provide estimation of these properties with good accuracy. It can be seen from Table 4.12, that the developed GC models for EUA_C, EUA_{NC}, ERA_C, ERA_{NC}, EFW_C, EFW_{NC}, ESW_C ESW_{NC}, ENS_C, ENS_{NC}, EAS_C, and EAS_{NC} have reasonable performance statistics. The estimation of these properties is based exclusively on the molecular structure of the chemical and allows the user to calculate Human Toxicity Potential (HTP) which is needed to perform life cycle impact assessment of the product (Rosenbaum, 2011) thus increasing the application range of the USEtox^{TM} model to a wide range of chemicals.

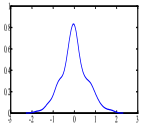
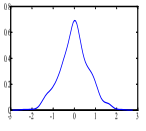
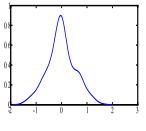
Table 4.12 Performance of MG method based property models for environmental-related properties analysed using step-wise regression method

Sl. no.	Property	L.H.S. of MG method based property prediction model	MG group-contribution model $f(X)=\sum_i N_i C_i + \sum_j M_j D_j + \sum_k E_k O_k$									
		$f(X)$	N	v	R^2	residual distribution plot	P_{rc} (± 1%)	P_{rc} (± 5%)	P_{rc} (± 10%)	SD	AAE	ARE [a]
1	Fathead minnow 96-hr LC_{50} (LC_{50} (FM)) in mol/lit	$-LogLC_{50}(FM) + FM_0$	809	541	0.78		8.53	31.52	54.02	0.69	0.48	21.56
2	*Daphnia magna* 48-hr LC_{50} (LC_{50} (DM)) in mol/lit	$-LogLC_{50}(DM) + DM_0$	320	124	0.82		16.25	39.06	62.50	0.74	0.49	16.16
3	Oral rat LD_{50} (LD_{50}) in mol/kg	$-LogLD50 - A_{LD50} - B_{LD50}MW$	5995	5617	0.73		1.52	6.92	13.61	0.43	0.35	16.40
4	Aqueous solubility ($LogW_S$) in gm/lit	$LogW_S - A_{W_S} - B_{W_S}MW$	4681	4311	0.78		3.12	14.36	28.63	0.99	0.73	----

5	Bioconcentra -tion factor (BCF)	LogBCF	662	423	0.78		8.91	19.49	30.82	0.63	0.47	----
6	Permissible exposure limit (OSHA-TWA) in mol/lit	-LogPEL	425	239	0.74		16.71	39.53	60.24	0.78	0.44	12.61
7	Photochemic al oxidation potential (PCO)	-LogPCO	639	488	0.83		6.42	16.9	26.30	0.22	0.13	8.37
8	Global warming potential (GWP)	LogGWP	51	31	0.87		15.69	37.25	56.86	0.41	0.29	11.57
9	Ozone depletion potential (ODP)	LogODP	28	12	0.89		17.86	21.4	28.5	0.30	0.16	----
10	Acidification potential (ODP)	LogAP	10	1	1.0		100.0	--	--	3.4E-04	2.1E-4	----

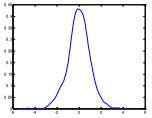
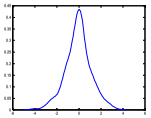
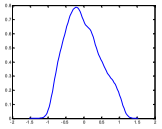
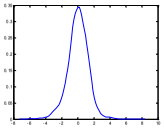
11	Emission to urban air (EUA_C) in cases/kg emitted (carcinogenic)	$-\text{Log}(EUA_C)+A_{EUA_C}$	456	214	0.70		16.23	40.13	63.60		0.70	0.50	10.61
12	Emission to urban air (EUA_{NC}) in cases/kg emitted (non-carcinogenic)	$-\text{Log}(EUA_{NC})+A_{EUA_{NC}}$	341	128	0.79		12.90	47.80	76.50		0.49	0.37	6.80
13	Emission to continental rural air (ERA_C) in cases/kg emitted (carcinogenic)	$-\text{Log}(ERA_C)+A_{ERA_C}$	470	229	0.75		15.74	39.15	64.89		0.67	0.51	8.88
14	Emission to continental rural air (ERA_{NC}) in cases/kg emitted (non-carcinogenic)	$-\text{Log}(ERA_{NC})+A_{ERA_{NC}}$	349	134	0.80		13.18	46.13	75.07		0.55	0.42	7.25

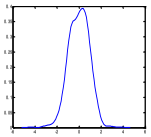
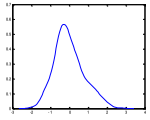
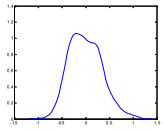
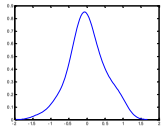
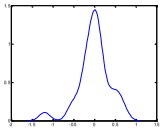
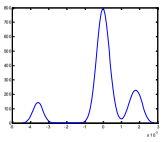
15	Emission to continental fresh water (EFW_C) in cases/kg emitted (carcinogenic)	$-Log(EFW_C)+A_{EFW_C}$	472	230	0.75		13.77	31.77	60.16	0.67	0.52	11.26
16	Emission to continental fresh water (EFW_{NC}) in cases/kg emitted (non-carcinogenic)	$-Log(EFW_{NC})+A_{EFW_{NC}}$	345	131	0.83		13.33	44.63	67.82	0.52	0.40	8.15
17	Emission to continental sea water (ESW_C) in cases/kg emitted (carcinogenic)	$-Log(ESW_C)+A_{ESW_C}$	477	235	0.81		15.30	37.94	67.71	0.79	0.61	8.69
18	Emission to continental sea water (ESW_{NC}) in cases/kg emitted (non-carcinogenic)	$-Log(ESW_{NC})+A_{ESW_{NC}}$	360	146	0.85		14.16	46.38	72.22	0.69	0.51	8.41
19	Emission to continental natural soil (ENS_C) in cases/kg emitted (carcinogenic)	$-Log(ENS_C)+A_{ENS_C}$	472	231	0.76		13.98	39.61	63.55	0.72	0.55	9.28

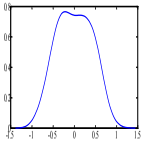
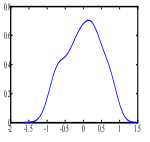
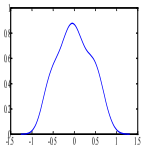
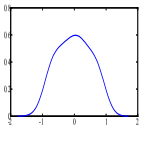
20	Emission to continental natural soil (ENS_{NC}) in cases/kg emitted (non-carcinogenic)	$-Log(ENS_{NC})+A_{ENS_{NC}}$	362	148	0.79		14.91	48.06	71.27	0.61	0.46	7.27
21	Emission to continental agri-cultural soil (EAS_C) in cases/kg emitted (carcinogenic)	$-Log(EAS_C)+A_{EAS_C}$	470	228	0.75		13.61	41.06	65.74	0.67	0.51	9.36
22	Emission to continental agri-cultural soil (EAS_{NC}) in cases/kg emitted (non-carcinogenic)	$-Log(EAS_{NC})+A_{EAS_{NC}}$	352	138	0.80		16.19	48.29	74.71	0.54	0.41	6.92

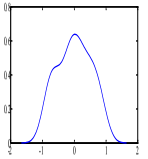
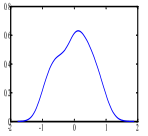
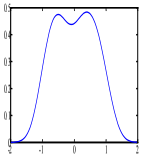
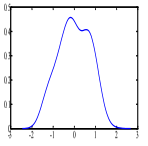
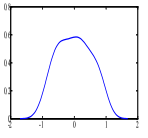
[a] ARE is not defined for LogWs, BCF, ODP and AP since these properties have both positive and negative values.

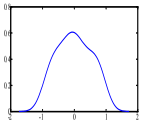
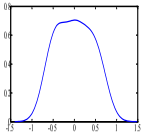
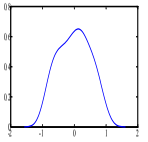
Table 4.13 Performance of CI method based property models for environmental-related properties

Sl. no.	Property	L.H.S. of CI method based property prediction model	**Atom connectivity index (CI) model** $f(X)=\sum_i a_i A_i + b\left({}^{v}\chi^{0}\right)+2c\left({}^{v}\chi^{1}\right)+d$									
		$f(X)$	N	v	R^2	residual distribution plot	P_{rc} (± 1%)	P_{rc} (± 5%)	P_{rc} (± 10%)	SD	AAE	ARE [a]
1	Fathead minnow 96-hr LC_{50} (LC_{50} (FM)) in mol/lit	$-\text{LogLC}_{50}(\text{FM})+\text{FM}_0$	809	796	0.56		3.96	16.70	34.20	0.98	0.75	40.47
2	*Daphnia magna* 48-hr LC_{50} (LC_{50} (DM)) in mol/lit	$-\text{LogLC}_{50}(\text{DM})+\text{DM}_0$	320	307	0.58		5.0	22.81	40.94	1.14	0.85	35.21
3	Oral rat LD_{50} (LD_{50}) in mol/kg	$-\text{LogLD50}-A_{LD50}-B_{LD50}MW$	5662	5647	0.60		1.02	5.35	11.48	0.48	0.40	18.49
4	Aqueous solubility ($LogW_S$) in gm/lit	$\log(W_S)-A_{W_S}-B_{W_S}MW$	4681	4676	0.62		2.22	9.98	19.80	1.29	0.98	----

5	Bioconcentra -tion factor (BCF)	LogBCF	662	648	0.53		1.66	6.19	12.54	0.92	0.74	----
6	Permissible exposure limit (PEL) in mol/lit	-LogPEL	411	397	0.64		4.87	16.79	33.09	0.78	0.61	20.10
7	Photochemic al oxidation potential (PCO)	-LogPCO	621	607	0.51		1.61	4.83	8.05	0.33	0.27	16.65
8	Global warming potential (GWP)	LogGWP	51	37	0.83		9.80	31.37	50.98	0.48	0.36	15.52
9	Ozone depletion potential (ODP)	LogODP	28	14	0.83		7.14	10.71	14.30	0.37	0.25	----
10	Acidification potential (ODP)	LogAP	10	1	1.00		70.0	100.0	--	0.0014	802E-04	--

11	Emission to urban air (EUA_C) in cases/kg emitted (carcinogenic)	$-\text{Log}(EUA_C)+A_{EUA_C}$	232	220	0.66		5.17	40.08	83.18	0.40	0.34	6.36
12	Emission to urban air (EUA_{NC}) in cases/kg emitted (non-carcinogenic)	$-\text{Log}(EUA_{NC})+A_{EUA_{NC}}$	259	247	0.66		7.72	40.92	69.11	0.49	0.41	7.50
13	Emission to continental rural air (ERA_C) in cases/kg emitted (carcinogenic)	$-\text{Log}(ERA_C)+A_{ERA_C}$	226	214	0.79		11.94	49.55	88.05	0.39	0.32	5.43
14	Emission to continental rural air (ERA_{NC}) in cases/kg emitted (non-carcinogenic)	$-\text{Log}(ERA_{NC})+A_{ERA_{NC}}$	257	245	0.74		7.78	39.69	68.48	0.53	0.44	7.57

15	Emission to continental fresh water (EFW_C) in cases/kg emitted (carcinogenic)	$-Log(EFW_C)+A_{EFW_C}$	286	274	0.65		7.34	36.36	61.53	0.52	0.44	8.51
16	Emission to continental fresh water (EFW_{NC}) in cases/kg emitted (non-carcinogenic)	$-Log(EFW_{NC})+A_{EFW_{NC}}$	259	247	0.70		9.26	33.59	60.61	0.54	0.44	9.02
17	Emission to continental sea water (ESW_C) in cases/kg emitted (carcinogenic)	$-Log(ESW_C)+A_{ESW_C}$	286	274	0.78		4.89	35.31	75.17	0.62	0.54	7.20
18	Emission to continental sea water (ESW_{NC}) in cases/kg emitted (non-carcinogenic)	$-Log(ESW_{NC})+A_{ESW_{NC}}$	291	279	0.77		5.84	32.30	65.63	0.72	0.60	8.76
19	Emission to continental natural soil (ENS_C) in cases/kg emitted (carcinogenic)	$-Log(ENS_C)+A_{ENS_C}$	285	273	0.61		6.66	38.59	74.38	0.52	0.44	6.89

20	Emission to continental natural soil (ENS_{NC}) in cases/kg emitted (non-carcinogenic)	$-Log(ENS_{NC})+A_{ENS_{NC}}$	247	235	0.70		9.31	39.67	72.06	0.53	0.45	7.08
21	Emission to continental agri-cultural soil (EAS_C) in cases/kg emitted (carcinogenic)	$-Log(EAS_C)+A_{EAS_C}$	240	228	0.68		8.33	42.50	88.33	0.42	0.36	5.76
22	Emission to continental agri-cultural soil (EAS_{NC}) in cases/kg emitted (non-carcinogenic)	$-Log(EAS_{NC})+A_{EAS_{NC}}$	247	235	0.70		8.50	40.89	74.08	0.49	0.42	6.94

[a] ARE is not defined for LogWs, BCF, ODP and AP since these properties have both positive and negative values.

The variables FM_0, DM_0, A_{LogWs}, B_{LogWs} A_{LD50}, B_{LD50}, A_{EUAC}, A_{EUANC}, A_{ERAC}, A_{ERANC}, A_{EFWC}, A_{EFWNC}, A_{ESWC}, A_{ESWNC}, A_{ENSC}, A_{ENSNC}, A_{EASC}, A_{EASNC} as defined in the functional forms, $f(X)$ given in Tables 4.12 and 4.13 are additional adjustable parameters of property prediction models. The values of these parameters together with the uncertainty in terms of 95% confidence interval are given in Table 4.14. The total list of groups and their contributions C_i, D_j, and E_k for the 22 environmental-related properties analysed in this work are given in Appendix B (see Tables B3-B4 for MG method based models analysed using step-wise regression method, and Tables B5-B7 for MG method based models analysed using simultaneous regression method). The list of atoms, their contributions a_i, adjustable model parameters (b and c), and the universal parameter d for CI method based property prediction models are also given in Appendix B (see Table B8).

Table 4.14 Values of universal constants (additional adjustable parameters)[a]

Universal constants	Step-wise method		Simultaneous method	
	Value	Uncertainty (95% CI)	Value	Uncertainty (95% CI)
FM_0	2.19	0.29	2.18	0.89
DM_0	2.97	0.84	3.59	1.94
A_{LogWs}	4.55	0.35	4.31	0.42
B_{LogWs}	0.34	0.05	0.34	0.11
A_{LD50}	1.94	0.09	1.94	0.08
B_{LD50}	0.002	0.02	0.002	0.01
A_{EUAC}	5.28	0.59	5.23	1.05
A_{EUANC}	6.82	2.35	7.07	2.59
A_{ERAC}	6.56	0.53	6.69	1.05
A_{ERANC}	7.55	3.08	9.53	5.06
A_{EFWC}	5.67	0.55	5.07	1.10
A_{EFWNC}	6.44	1.97	7.33	2.86
A_{ESWC}	8.40	0.62	9.33	1.15
A_{ESWNC}	8.64	4.16	10.07	5.60
A_{ENSC}	6.48	0.56	5.93	1.04
A_{ENSNC}	7.03	2.55	6.42	1.94
A_{EASC}	6.29	0.54	5.49	1.01
A_{EASNC}	6.97	2.50	6.06	1.59

[a] values of universal constants for the CI models are the same as those based on the step-wise method.

4.2.4 Selection of quantity of property data for parameter estimation

As illustrated in section 4.1 of this chapter, by including all of the available experimental data of the property in the regression it is possible to improve the prediction accuracy and application range of the property model. Therefore, all of the available experimental data of properties of pure components are considered for modeling environmental-related properties. To further illustrate this point, analysis of property model for oral rat LD_{50}, fathead minnow 96-h LC_{50}, and emission to continental rural air (carcinogenic (ERA_C) and non-carcinogenic (ERA_{NC})) is presented here. The whole experimental data-sets of these properties (5995 data-points for oral rat LD_{50}, 809 data-points for fathead minnow 96-hr LC_{50}, 456 data-points for emission to continental rural air (carcinogenic (ERA_C), and 341 data-points for emission to continental rural air (non-carcinogenic (ERA_{NC}))) is divided randomly in 5 subsets (A, B, C, D, and E) of equal size. The property model is trained on 4 subsets (using simultaneous regression method) and 1 subset is used for testing purpose. This procedure is repeated 5 times so that all subsets are used for testing purpose. The results in terms of SD, AAE, and ARE for training sets and for test sets is presented in Table 4.15(a) for oral rat LD_{50}, Table 4.15(b) for fathead minnow 96-h LC_{50}, Table 4.15(c) for emission to continental rural air (carcinogenic (ERA_C) and in Table 4.15(d) for emission to continental rural air (non-carcinogenic (ERA_{NC})). The *MSECV*, which is mean squared error of cross-validation (Mevik and Cederkvist, 2004) calculated using Eq. (4.3) is also given in Tables 4.15(a)-4.15(d).

$$MSECV = \frac{1}{N_L}\sum_{k=1}^{K}\sum_{j\in L_k}\left(X_j^{exp} - X_j^{pred}\right)^2 \tag{4.3}$$

Where, N_L is the number of data-points in the training set, K = number of subsets (5 in this analysis), and L_k is the number of data-points in the subsets.

From Table 4.15(a), comparison of model performance for training sets and test sets show that the predictive capability of the model for oral rat LD_{50} is fairly good. This is mainly due to the large amount of available experimental data of oral rat LD_{50} for the training purpose. For test sets, if we compare the SD, AAE, and ARE values calculated using the parameters obtained by regressing training set with those that are calculated using the parameters obtained by regression of the whole data-set, we find that better model performance statistics (lower SD, lower AAE, and lower ARE) is obtained when we use model parameters that are estimated using all of the experimental data-points in the regression.

For fathead minnow 96-hr LC_{50}, emission to continental rural air (carcinogenic (ERA_C)), and emission to continental rural air (non-carcinogenic (ERA_{NC})) it can be seen from Tables 4.15(b)-4.15(d) that the model performance for test sets is poor as compared to those of training sets and this is due to the small amount of available experimental data of these

properties for the training purpose. For these properties, it can be observed that the SD, AAE, and ARE values for test sets calculated using the model parameters as obtained by regression of the whole data-set are much better than those that are calculated using the parameters estimated using the training set indicating the importance of considering all of the available experimental data-points for the regression purpose. To sum up, this analysis shows the robustness of the approach for the development of property models for the estimation of pure component properties.

Table 4.15(a) Performance of model for Oral Rat LD_{50} based on different combinations of training sets and test sets

datasets used for		model performance statistics for training set				model performance statistics for test set using the parameters estimated from regression of the training set			model performance statistics for test set using the parameters estimated from regression of the whole data-set (containing 5995 data-points)		
training purpose	testing purpose	*MSECV*	SD Log(mol/kg)	AAE Log(mol/kg)	ARE %	SD Log(mol/kg)	AAE Log(mol/kg)	ARE %	SD Log(mol/kg)	AAE Log(mol/kg)	ARE %
A, B, C, D	E	0.1812	0.4257	0.3479	15.97	0.4628	0.3732	17.32	0.4220	0.3424	15.91
A, B, C, E	D	0.1796	0.4238	0.3456	15.90	0.4755	0.3839	17.72	0.4287	0.3506	16.21
A, B, D, E	C	0.1805	0.4248	0.3462	15.97	0.4754	0.3823	17.23	0.4251	0.3500	15.90
A, C, D, E	B	0.1788	0.4229	0.3449	15.89	0.4677	0.3813	17.20	0.4338	0.3536	16.00
B, C, D, E	A	0.1794	0.4236	0.3455	15.86	0.4694	0.3848	17.93	0.4302	0.3532	16.46
average performance		0.1799	0.4241	0.3460	15.91	0.4701	0.3811	17.48	0.4279	0.3499	16.09

Table 4.15(b) Performance of model for fathead minnow 96-h LC_{50} based on different combinations of training sets and test sets

datasets used for		model performance statistics for training set				model performance statistics for test set using the parameters estimated from regression of the training set			model performance statistics for test set using the parameters estimated from regression of the whole data-set (containing 809 data-points)		
training purpose	testing purpose	*MSECV*	SD Log(mol/kg)	AAE Log(mol/kg)	ARE %	SD Log(mol/kg)	AAE Log(mol/kg)	ARE %	SD Log(mol/kg)	AAE Log(mol/kg)	ARE %
A, B, C, D	E	0.3400	0.5831	0.4015	19.68	1.3753	0.8615	27.04	0.6732	0.4786	15.68
A, B, C, E	D	0.3339	0.5778	0.3991	19.10	1.3944	0.9325	28.91	0.6854	0.4778	17.65
A, B, D, E	C	0.3624	0.6020	0.4237	14.99	1.3517	0.9127	47.19	0.6581	0.4802	33.19
A, C, D, E	B	0.3645	0.6037	0.4201	20.39	1.4857	0.9072	28.26	0.6399	0.4654	14.62
B, C, D, E	A	0.3453	0.5876	0.4142	17.62	1.5178	0.9710	35.12	0.6722	0.4831	21.19
average performance		0.3492	0.5908	0.4117	18.35	1.4249	0.9169	33.30	0.6657	0.4770	20.47

Table 4.15(c) Model performance for emission to Urban Air (Carcinogenic) based on different combinations of training and test sets

datasets used for		model performance statistics for training set				model performance statistics for test set using the parameters estimated from regression of the training set			model performance statistics for test set using parameters estimated from regression of the whole data-set (containing 456 data-points)		
training purpose	testing purpose	*MSECV*	SD Cases/kg	AAE Cases/kg	ARE %	SD Cases/kg	AAE Cases/kg	ARE %	SD Cases/kg	AAE Cases/kg	ARE %
A, B, C, D	E	0.2050	0.4528	0.3024	5.86	1.8293	1.1974	27.82	0.5386	0.4135	9.16
A, B, C, E	D	0.2206	0.4697	0.3252	6.55	1.4664	0.9823	17.63	0.4766	0.3534	6.60
A, B, D, E	C	0.1675	0.4093	0.2871	5.52	1.6849	1.2268	25.79	0.6160	0.4165	9.52
A, C, D, E	B	0.2111	0.4595	0.3052	6.42	2.0921	1.3187	24.12	0.4854	0.3480	6.33
B, C, D, E	A	0.2182	0.4671	0.3267	6.63	1.3597	1.0115	19.55	0.4572	0.3478	6.55
average performance		0.2045	0.4517	0.3093	6.19	1.6865	1.1473	22.98	0.4713	0.3479	6.44

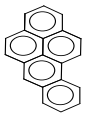
Table 4.15(d) Model performance for emission to Urban Air (Non-Carcinogenic) based on different combinations of training and test sets

datasets used for		model performance statistics for training set				model performance statistics for test set using the parameters estimated from regression of the training set			model performance statistics for test set using parameters estimated from regression of the whole data-set (containing 341 data-points)		
training purpose	testing purpose	*MSECV*	SD Cases/kg	AAE Cases/kg	ARE %	SD Cases/kg	AAE Cases/kg	ARE %	SD Cases/kg	AAE Cases/kg	ARE %
A, B, C, D	E	0.0655	0.2560	0.1716	3.14	2.3797	1.4854	26.38	0.3872	0.2801	5.03
A, B, C, E	D	0.0997	0.3157	0.2217	4.12	1.7248	1.0980	19.22	0.3693	0.2708	4.89
A, B, D, E	C	0.0831	0.2882	0.1894	3.50	2.4662	1.5963	30.72	0.3684	0.2615	5.01
A, C, D, E	B	0.0846	0.2909	0.1936	3.51	3.0908	1.6401	29.28	0.3980	0.3007	5.62
B, C, D, E	A	0.1097	0.3313	0.2289	4.23	8.7061	3.9458	80.85	0.3096	0.2166	3.80
average performance		0.0885	0.2964	0.2010	3.69	3.6735	1.9531	37.28	0.3665	0.2659	4.87

4.2.5 Application Example

The application of the developed property models for the estimation of environmental-related properties and quantification of uncertainties of predicted property values is illustrated by considering predictions of $LogW_s$ (using model parameters obtained from simultaneous regression method) for the chemical, Benzo[a]pyrene, (CAS No. 50-32-8) which is a polycyclic aromatic hydrocarbon and is highly carcinogenic. The experimentally measured value of $LogW_s$ (Log mg/L) for Benzo[a]pyrene is -2.79. Table 4.16 provides information of first-order, second-order, and third-order MG groups used to represent Benzo[a]pyrene, their frequency (that is, occurrences in the structure) and the contributions for each group ($LogW_{s1i}$, $LogW_{s2j}$, and $LogW_{s3k}$) taken from Appendix B (given in Tables B5-B7). Using this information and the universal constants of the property model for $LogW_s$, the value of $LogW_s$ for Benzo[a]pyrene was estimated as -2.64 (with absolute error = |-2.79 – (-2.64)| = 0.15).

Table 4.16 Estimation of $LogW_s$ of Benzo[a]pyrene

Benzo[a]pyrene	molecular structure	
CAS No. 50-32-8		
molecular formula: $C_{20}H_{12}$		
first-order groups	**occurrences**	**contribution**
aCH	12	-4.5565
aC fused with aromatic ring	8	-4.7557
second-order groups	**occurrences**	**contribution**
No second-order groups are involved		
third-order groups	**occurrences**	**contribution**
AROM.FUSED[2]	1	-0.0759
AROM.FUSED[3]	1	-0.1255
AROM.FUSED[4p]	2	0.0500

$$LogW_s = A_{LogWs} + (B_{LogWs} \cdot MW) + \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k E_k O_k = -2.64$$

Note: estimated value of $LogW_s$ using Marrero and Gani (2002) method = -2.20

As a next step, the uncertainty of the estimated $LogW_s$ is quantified. For this purpose, information of covariance $COV(\boldsymbol{P}^*)$ of the involved groups and the universal constants A_{LogWs} and B_{LogWs} and also the local sensitivity $J(\boldsymbol{P}^*)$ of the $LogW_s$ model is needed. The covariance of the involved groups (as listed in Table 4.16) and universal constants A_{LogWs} and B_{LogWs} was noted from the overall covariance matrix for all the groups of the $LogW_s$ model analysed using simultaneous regression method. In Table 4.17, only lower triangular elements are shown since the upper triangular matrix elements are identical to the lower ones. Table 4.18 lists the local sensitivity of the $LogW_s$ model with respect to the model parameters (for contributions listed in Table 4.16 and universal constants A_{LogWs} and B_{LogWs}).

Table 4.17 Covariance matrix $COV(P^*)$ with dimensions (7×7)

	A_{LogWs}	B_{LogWs}	aCH	aC	AROM.FUSED[2]	AROM.FUSED[3]	AROM.FUSED[4p]
A_{LogWs}	0.0154						
B_{LogWs}	-1.28E-07	4.97E-07					
aCH	-0.0025	-8.1E-06	6.71E-04				
aC	-0.002	-4.7E-06	-3.7E-04	0.0048			
AROM.FUSED[2]	7.7E-05	-4.3E-06	-5.9E-04	-0.0047	0.0113		
AROM.FUSED[3]	-0.0013	-2.4E-06	-8.9E-06	-0.0084	0.0111	0.0375	
AROM.FUSED[4p]	-4.5E-04	1.6E-06	8.25E-06	-0.0092	0.009	0.0136	0.0283

Table 4.18 Local sensitivity $J(P^*)$ with dimensions (1×7) of $LogW_s$ model with respect to the model parameters

δLogWs/ δA_{LogWs}	*δLogWs /* δB_{LogWs}	*δLogWs /* /δaCH	*δLogWs /* /δaC	*δLogWs /* δ AROM.FUSED[2]	*δLogWs /* δ AROM.FUSED[3]	*δLogWs /* δ AROM.FUSED[4p]
1.0	252.31	12	8	1	1	2

To calculate the confidence intervals of estimated property values, say the 95% confidence intervals of the estimated $LogW_s$ value, the covariance matrix $COV(\boldsymbol{P}^*)$ given in Table 4.17 and the local sensitivity $J(\boldsymbol{P}^*)$ given in Table 4.18 are substituted in Eq. (3.13). For 95% confidence interval calculation, the t-distribution value corresponding to 0.05 /2 percentile (i.e. $\alpha_t/2$ percentile) and with 4311 degrees of freedom (taken from Table 4.12) is obtained by solving Eq. (3.12) for t and this value is 1.9604. The predicted value of the $LogW_s$ is -2.64 (see Table 4.16). The calculated 95% confidence intervals of the estimated $LogW_s$ value is therefore,

$$\mathrm{LogW_s}^{\,pred}_{(1-0.05)} = \underbrace{\mathrm{LogW_s}^{\,pred}}_{-2.64} \pm \underbrace{\sqrt{\mathrm{diag}\left(J\left(\boldsymbol{P}^*\right)COV\left(\boldsymbol{P}^*\right)J\left(\boldsymbol{P}^*\right)^T\right)}}_{0.2134} \cdot \underbrace{t\left(\nu,\alpha_t/2\right)}_{1.9604} = -2.64 \pm 0.41$$

It can be observed that the experimental value of the $LogW_s$ (-2.79) lies within the predicted confidence intervals indicating reliability of the developed model for estimating property values of $LogW_s$ and uncertainties of the estimated values. This is further illustrated in Figure 4.5 by plotting the experimental values of $LogW_s$ and the calculated 95% confidence intervals (shown as vertical bars) for the entire experimental data-set of $LogW_s$ used for the regression purpose. About 42% of the experimental values in the data-set (with 4681 data points) of $LogW_s$ falls within the confidence intervals calculated at 95% confidence level.

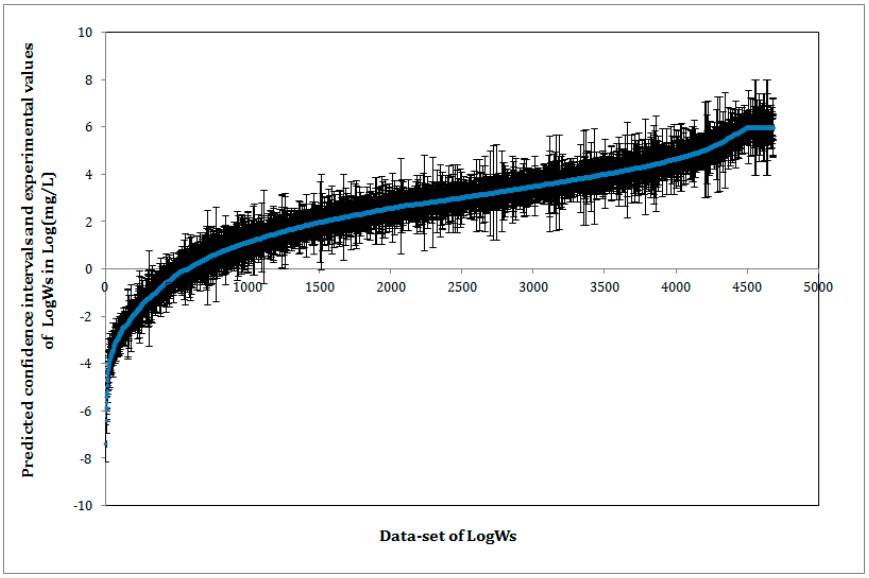


Figure 4.5 Experimental values of $LogW_s$ and the calculated 95% confidence intervals versus data-set of $LogW_s$

4.2.6 Summary

Property models for environmental-related properties based on the GC^+ approach have been developed with the objective of providing reliable estimation of these properties together with the uncertainties of the estimated values. The estimation of environmental-related properties using these models requires only the molecular structure of pure components. Experimental data-sets of environmental-related properties taken from the database of US Environmental Protection Agency (EPA) and from the database of USEtox^{TM} are used for the regression purpose. In total, 22 environmental-related properties of organic chemicals have been modeled and analysed. The application of the developed property models to estimate environmental-related properties and the uncertainties of the estimated property values is

illustrated through an application example. The developed property models provide reliable estimates of environmental-related properties needed to perform design and analysis of sustainable processes and allow one to evaluate the effect of uncertainties of estimated property values on the quality and reliability of the design of sustainable processes.

4.3 Further Improvement of GC Models using 'Molecular Structural Similarity Criteria' Approach

In the following section, the application of method based on the 'molecular structural similarity criteria' based approach is illustrated through analysis of different classes of properties such as $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, T_c using MG method based property models. A method to find the minimum data-set for the parameter regression as discussed in section 3.2.1 of chapter 3 is applied for assessing the reliability of predictions from the improved GC model for $\Delta_f H^o_{gas}$.

4.3.1 Development of a GC model for $\Delta_f H^o_{gas}$ with *chemical accuracy*

To understand the reactivities and chemical equilibria, knowledge of $\Delta_f H^o_{gas}$ of pure components is necessary. This knowledge is of very broad interest and includes both experimental chemistry and process technology. Experimental measurements of $\Delta_f H^o_{gas}$ are one of the ways of collecting reliable and accurate data. However, the limitations and challenges including, the determination of $\Delta_f H^o_{gas}$ of unstable species, the required purity of samples, time and cost of experiments etc. faced by experimentalists are well-known. In such a situation, the most convenient and practical approach has been to employ property prediction models to estimate $\Delta_f H^o_{gas}$ of components from their molecular structures. In this context, two important classes of property models are widely used: *ab initio* quantum mechanics based property models and the GC based property models. A review by Van Speybroeck et al. (2010) provides the state-of-the-art of estimation of $\Delta_f H^o_{gas}$ using these methods. Various GC methods for the prediction of $\Delta_f H^o_{gas}$ of pure components have been developed by many authors including Joback and Reid (1987), Constantinou and Gani (1994), Marrero and Gani (2001), Benson and co-workers (1969, 1993), Domalski and Hearing (1993) to name a few.

The "holy grail" in the field of computational thermochemistry is to arrive at *chemical accuracy*. Peterson et al. (2012) recently recalled that "in the thermochemistry literature this is almost universally interpreted as 1 kcal/mol or about 4 kJ/mol. However, a MUD (mean unsigned deviation) of 1 kcal/mol does not actually represent *chemical accuracy* analogous to experiment since for experimentalists the latter is generally based on the 95% confidence level or about 2 standard deviations." Consequently, a model with the more stringent MUD of

0.5 kcal/mol or about 2 kJ/mol is required as a reliable and predictive tool for the experimentalists. Currently, there is no GC method (neither any other available method) capable of achieving *chemical accuracy* of ±2 kJ/mol for a broad range of organic molecules. Therefore, in this work, a GC model based on the MG method has been developed which achieves this accuracy. To achieve the accuracy of ± 2 kJ/mol, 'molecular structural similarity criteria' based approach which allows efficient use of knowledge of the experimental data of $\Delta_f H^o_{gas}$ and the molecular structural information is employed. Using this approach, new structural parameters (that is, functional groups) are defined and included in the MG model to provide additional structural information for components having large correlation errors and to thereby improve the accuracy of $\Delta_f H^o_{gas}$ predictions through better correlation of data. The developed property model for $\Delta_f H^o_{gas}$ is based on simple principles and is computational undemanding so it can be used 'on the fly' by the chemist or process engineer.

As noted in the above, no current GC method achieves *chemical accuracy* of ±2 kJ/mol and hence it is required to improve on specific points in order to arrive at the desired accuracy. The previous MG model for $\Delta_f H^o_{gas}$ was developed using 787 data-points with an average absolute deviation (AAD) of 7.25 kJ/mol (Van Speybroeck et al., 2010). The GC model developed in this work (see section 4.2 of this chapter) has AAD of 4.15 kJ/mol. Therefore, to develop an even better model that can provide the desired *chemical accuracy*, many key issues are addressed through following steps.

4.3.1.1 Database

In the present work, a data-set containing 861 experimentally measured values for a wide variety of components (hydrocarbons, oxygen containing, nitrogen containing, multi-functional components, etc.) is used for the regression purpose. These experimental values of $\Delta_f H^o_{gas}$ have been taken from the extended CAPEC database (2001) and an extended set of nitrogen-containing molecules to allow for an optimal description of the large variety of nitrogen containing chemical species obtained from a member company of CAPEC-industrial consortium. The description of this data-set based on various classes of organic components is given in Table 4.19.

Table 4.19 Class-wise description of the data-set of $\Delta_f H^o_{gas}$

Class[a]	CH	O	N	S	Cl	F	Br	I	P	Multi
N	264	257	160	27	30	5	9	5	1	103

[a] CH, Hydrocarbons; O, Oxygenated; N, Nitrogenated; S, Sulphur containing; Cl, Chlorinated; F, Fluorinated; Br, Brominated; I, Iodinated; P, Phosphorous containing; Multi, Multifunctional components.

The experimental values of $\Delta_f H^o_{gas}$ are plotted as a function of increasing carbon number of components belonging to different classes, such as, *n*-alkanes, *n*-alkenes, cyclic alkanes, mercaptans, primary alcohols, nitriles, mono-aromatic hydrocarbons, polycyclic aromatics, carboxylic acids, ketones, nitriles, and esters (see Figure 4.6(a)). For each class, except for cyclic alkanes, the relationship between the property as a function of the carbon number is established. The function is found to be linear in nature (confirming the principle of GC approach), which makes it possible to predict $\Delta_f H^o_{gas}$ of other components from these classes of components with good accuracy.

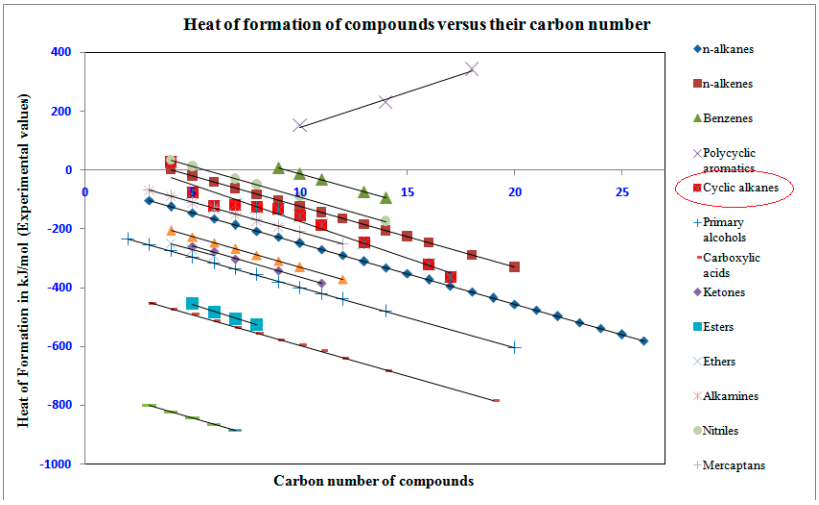


Figure 4.6(a) Plot of $\Delta_f H^o_{gas}$ versus carbon number of pure components.

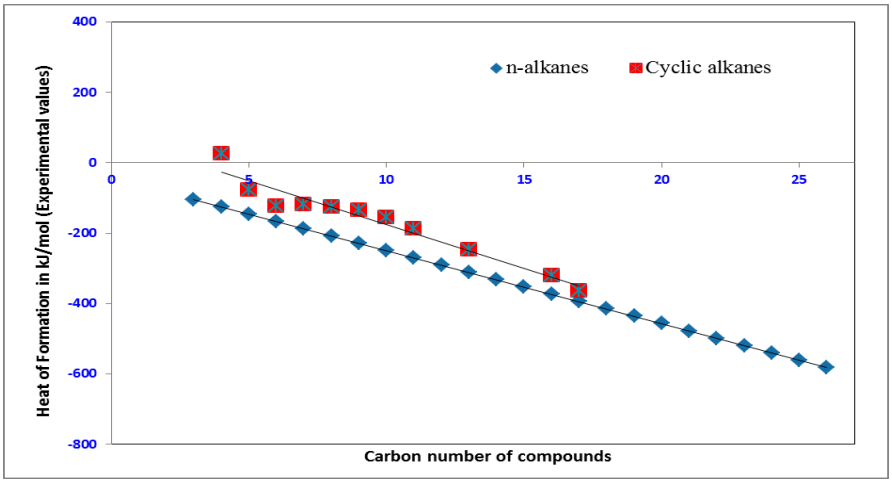


Figure 4.6(b) Plot of $\Delta_f H^o_{gas}$ of cyclic alkanes and *n*-alkanes versus carbon number

For cyclic alkanes (see Figure 4.6(b)), a plot of $\Delta_f H^o_{gas}$ versus their carbon number is not linear in nature but jagged owing to the strain energy of the ring (Figure 4.6). The main reason, as pointed out by Cohen and Benson (1993), is that the CH_2 groups in one cyclic alkane (for example, in cyclohexane) behaves differently than that present in another cyclic alkane (for example, in cyclopentane or in cycloheptane) making it difficult for the same group-parameter to predict $\Delta_f H^o_{gas}$ for cyclic alkanes with good accuracy. For accurate modeling of $\Delta_f H^o_{gas}$ of cyclic alkanes, one needs to incorporate in the GC model, a special correction term depending on the type of the ring of cyclic alkanes. In the case of polycyclic aromatics, the uncertainty in the measurement of $\Delta_f H^o_{gas}$ seems to be significant. For example, various values of $\Delta_f H^o_{gas}$ of anthracene are reported in the literature ranging from 209.1 kJ/mol (as given by Parks et al., 1946) to 229.4 kJ/mol (as given by Roux et al., 2008). In such a situation, selection of critically evaluated and recommended data is very important. A similar issue is associated with the $\Delta_f H^o_{gas}$ of naphthacene with one value reported as 331.6 ± 4.4 kJ/mol by Nagano (2002) and another reported value being 342.6 ± 5.9 kJ/mol by Roux et al. (2008). In this work, the values of $\Delta_f H^o_{gas}$ of polycyclic aromatics critically evaluated and recommended by Roux et al. (2008) are considered. Also, the experimental values of $\Delta_f H^o_{gas}$ of components in the data-set were checked with the values available in NIST® database (Frenkel et al., 2005) to avoid further inconsistencies in the data-set.

4.3.1.2 Selection/Verification of most appropriate model function

Figure 4.6 shows that $\Delta_f H^o_{gas}$ can be modeled using a linear form of the function. For $\Delta_f H^o_{gas}$, the definition of $f(X)$ used in the MG method is, $(\Delta_f H^o_{gas} - H_{f0})$. Here, the variable H_{f0} is an additional adjustable parameter (universal constant) of the property model.

4.3.1.3 Apply "molecular structural similarity criteria" based approach

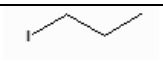
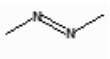
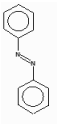
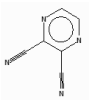
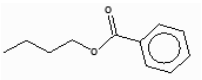
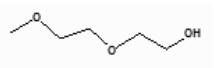
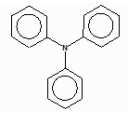
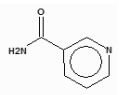
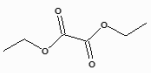
The MG method is first applied to components present in the data-set using developed model parameters (see section 4.1 of this chapter) to compute $\Delta_f H^o_{gas}$ values and to obtain their deviation from the experimental values. Based on these observed deviations, components in the data-set are classified in terms of absolute deviation ranges, such as, 0-1, 1-4, 4-6, etc. (see Table 4.20). Components having absolute deviation of > 6 kJ/mol are identified and considered for further analysis (146 such components are found). These are mainly components with N=N fragment, polycyclic aromatics containing N atom, cyclic alkanes, components containing ring structures, quinolines, iodine containing components, and glycols. Hence, in the present work, additional experimental data of $\Delta_f H^o_{gas}$ of N atom containing components are included in the data-set to better describe the large variety in components involving N atom.

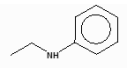
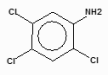
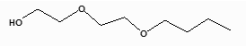
Table 4.20 Classification of data-set of $\Delta_f H^o_{gas}$ based on the range of observed deviation

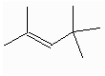
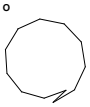
Range of absolute deviation in kJ/mol	Number of components from the data-set	Main types of components observed in the deviation range
0 – 1	334	*n*-alkanes, *n*-alkenes, Carboxylic acids, Esters, Nitriles, Primary alcohols, Halogenated hydrocarbons.
1 – 4	289	Ketones, Mercaptans, Fatty acids, Aromatic-OH containing components, Fatty alcohols.
4 – 6	92	Cycloalkenes, Polycyclic aromatics, Aliphatic components containing chlorine and bromine, Pyridine structure containing components, Ethers.
6 – 10	64	Cyclolalkanes, components containing ring structures, Quinolines, Glycols.
> 10	82	Polycyclic aromatics with N atom, N=N containing components, Compounds containing Iodine and Bromine, Compounds with ring structures.

The molecular structures of the identified 146 components (having absolute deviation of > 6 kJ/mol) are now analysed and new structural third-order groups are defined and included in the MG method based GC model as additional new third-order groups, providing more structural information. A list of these new third-order groups along with sample assignments are given in Table 4.21.

Table 4.21 List of new third-order groups defined for the Marrero and Gani GC-method with sample assignments

Sl. No.	New third-order group	Sample assignment [a]	Molecular structure of sample assignment
1	$CH_3 – (CH_2)_m –I$ m in 0,1,2…m	Propane, 1-iodo- (1)	
2	$aC– (N_{cyc} =CH_{cyc}) –S_{cyc}$	Benzothiazole (1)	
3	$aC– (N_{cyc} =CH_{cyc}) –O_{cyc}$	Benzoxazole (1)	
4	$CH_3– (CH_2)_m– (N=N)$m in 0,1,2…m	Azomethane (1)	
5	aC– (N=N)–aC	Cis-Azobenzene (1)	
6	aN– aC –CN	2,3-dicyanopyrazine (2)	
7	$CH_3 – (CH_2)_m –COO–aC$	Benzoic acid, butyl ester (1)	
8	$CH_3O –(CH_2)_m –O\ CH_2\ CH_2OH$	Diethylene glycol methyl ether (1)	
9	$(aN–aCH)_3$	1,3,5-Triazine (1)	
10	aC– aN –aC	Triphenylamine (1)	
11	$aC–CONH_2–aC–aN$	Nicotinamide	
12	$CH_3 –CH_2COO$ m in 0,1,2…m	Ethanedioic acid, diethyl ester (2)	

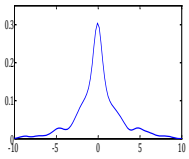
13	aC–aN–aC–$(CH_3)_m$ or $(OH)_m$ or $(Cl)_m$ m in 0,1	2-Methylquinoline (1)	
14	OH–$(aC)_2$–aN	3-Hydroxypyridine (1)	
15	aC–N_{cyc} –$(CH_{cyc}$=$CH_{cyc})$	1-Methylindole (1)	
16	Pyrazine in fused rings	Phenazine (1)	
17	CH_3 – $(CH_2)_m$ – CH_2COO m in 1,2…m	Decanoic acid, methyl ester (1)	
18	CH_3 – $(CH_2)_m$ –aC–NH m in 0,1,2…m	Benzenamine, N-ethyl- (1)	
19	aC– (aC– CH_3) in fused rings	Naphthalene, 1-methyl (1)	
20	$(aC\text{-}Cl)_2$– aCH– aC-Cl	2,4,5-trichloroaniline (1)	
21	CH_3 – $(CH_2)_m$ – CH_2O m in 1,2…m	Ethanol, 2-(2-butoxyethoxy) - (1)	
22	Cl–CH-Cl	1,1-Dichloropropane (1)	
23	NH_{cyc} –$(CH_{cyc})_m$ m in 3,4…m	Hexamethyleneimine (1)	
24	CH_3–N_{cyc}–CO_{cyc}	1-Methyl-2-pyrrolidone (1)	
25	aC-NH_{cyc} in different rings	2,3-Dihydroindole (1)	
26	5 member ring	1H-Pyrazole (1)	
27	6 member ring	Piperazine (1)	

28	7 member ring	Cycloheptanone (1)	
29	$(CH_3)_3C-(CH_2)_m-(CH_3)_3C$ m in 1,2…m	2,2,4,4-Tetramethylpentane (1)	
30	$(CH_3)_3C-CH=C$	2-Pentene, 2,4,4-trimethyl- (1)	
31	$(C_n H_{2n})$n in 4,5,6	Cyclobutane (1)	
32	$(C_n H_{2n})$ n in 7,8,9	Cycloheptane (1)	
33	$(C_n H_{2n})$ n in 10, 11,….n	Cyclodecane (1)	
34	(CH_{cyc}) -CO_{cyc}-(CH_{cyc})	Cyclododecanone (1)	

[a] The number in bracket indicates the occurrence of a particular group in the molecular structure of the sample.

Using the data-set of $\Delta_f H^o_{gas}$ and the GC model (with the new functional groups), parameter estimation and uncertainty analysis is performed to estimate the GC model parameters and to quantify the uncertainties (95% confidence intervals) of predicted $\Delta_f H^o_{gas}$ values. The performance statistics for the developed model for $\Delta_f H^o_{gas}$ is summarized in Table 4.22.

Table 4.22 Model performance statistics for the developed GC model for $\Delta_f H^o_{gas}$

Data-points	Correlation coefficient	Residual Distribution plot	SD kJ/mol	AAE kJ/mol	AE_{max} kJ/mol
861	0.9999		2.61	1.75	9.82

The residuals for data-points considered in the parameter estimation are plotted in the form of a residual distribution plot and this plot follows a normal distribution curve suggesting that the assumption of normal distribution of errors is valid for the employed parameter estimation error method. In Table 4.22, AE_{max} is the maximum value of the deviation found in the regression step. It can be seen from Table 4.22 that the experimental data have been fitted to a very good accuracy with AAE of 1.75 kJ/mol (within the required *chemical accuracy* of ±2 kJ/mol) and a low value of AE_{max} of 9.82 kJ/mol indicating the reliability of the developed model for predicting $\Delta_f H^o_{gas}$ for a wide variety of organic components involving C, H, O, N, S, Cl, Br, F, I and P atoms. The total list of first-order, second-order and third-order groups and their contributions C_i, D_j, and E_k are given in Appendix C (see Tables C1-C3). The model performance statistics based on individual classes of components is given in Table 4.23. It can be observed that the developed model for $\Delta_f H^o_{gas}$ performs very well for all types of components. It can also be seen that the inclusion of new experimental data of components containing N atom in the analysis has helped to achieve better prediction performance for these components. However, for organic components containing F, Br, I, and P atoms, additional experimental data are necessary to verify the prediction performance of the developed GC model.

Table 4.23 Performance of improved GC model for $\Delta_f H^o_{gas}$ for various classes of components

Class of components	Data-points	SD kJ/mol	AAD kJ/mol	AE_{max} kJ/mol
Hydrocarbons	264	1.84	1.68	9.82
O Containing	257	1.85	1.72	9.73
N Containing	160	1.80	1.71	7.43
S Containing	27	0.95	0.76	3.74
Cl Containing	30	2.06	2.01	7.26
F Containing	5	1.04	0.47	2.33
Br Containing	9	2.80	3.77	9.01
I Containing	5	1.21	1.11	2.78
P Containing	1 [a]	0	0	0
Multifunctional	103	2.25	2.01	8.16

[a] The number single functional components containing P atom in the data-set is 1. However, there are multifunctional compounds containing P atom in the data-set.

The distribution of the data-points based on the range of the observed absolute deviation is given in Table 4.24. There are now only 39 components in the data-set (as compared to 146 components, see Table 4.20) for which the deviation is greater than 6 kJ/mol but less than 10

kJ/mol. Note that out of these 39 components, 15 did not get any new groups, and for 24 components although improvements in the estimation of $\Delta_f H^o_{gas}$ are observed, the absolute deviations are still greater than 6 kJ/mol (but less than 10 kJ/ mol). For the 15 components that did not get any new groups, there were simply not enough components with similar structures to justify the addition of new groups. In Table 4.23, the AE_{max} of 9.82 kJ/mol corresponds to Cyclononane (a cyclic ring hydrocarbon), 9.73 kJ/mol corresponds to dimethyl ether (an ether), 7.43 kJ/mol corresponds to 1,2,3,4-Tetrahydroquinoline (aromatic component containing N atom), 9.01 kJ/mol corresponds to Butane, 2-bromo- (a Br containing component), and 8.16 kJ/mol corresponds to 8-Hydroxy-2-methylquinoline (a multifunctional component).

Table 4.24 Distribution of data-points based on the range of absolute deviations from the developed GC model for $\Delta_f H^o_{gas}$

Class of components	data-points	(0-1) kJ/ mol	(1-4) kJ/ mol	(4-6) kJ/ mol	> 6 and < 10 kJ/ mol
Hydrocarbons	264	123	117	10	14
O Containing	257	123	102	24	8
N Containing	160	73	63	21	3
S Containing	27	18	9	0	0
Cl Containing	30	11	12	5	2
F Containing	5	4	1	0	0
Br Containing	9	2	3	3	1
I Containing	5	3	2	0	0
P Containing	1	1	0	0	0
Multifunctional	103	51	29	12	11

4.3.1.4 Comparison of prediction performance of the developed GC model with currently used prediction methods

The developed GC model for $\Delta_f H^o_{gas}$ is compared with the Benson's group-additivity method (1993) which is well-known method for the estimation of $\Delta_f H^o_{gas}$ of a wide range of organic components. For the purpose of comparison, a common data-set of 799 components that can be described by the developed GC model in this work as well as by the Benson's method (1993) is considered. Table 4.25 provides the comparison of the prediction performance for each class of organic components (estimation of $\Delta_f H^o_{gas}$ based on the Benson's method (1993) is performed using the NIST® Structures and Properties software program). Table

4.26 provides comparison of the distribution of data-points based on the range of absolute deviations from the developed GC model and from the Benson's method.

Table 4.25 Comparison of prediction performance between the improved GC model and the Benson's method (1993)[a]

Class of components	Data-points	SD kJ/mol		AAE kJ/mol		AE_{max} kJ/mol	
		This work	Benson's method	This work	Benson's method	This work	Benson's method
Hydrocarbons	259	1.74	5.58	1.61	3.00	9.82	56.60
O Containing	256	1.84	5.44	1.73	3.75	9.73	42.00
N Containing	135	1.64	8.10	1.55	5.04	7.43	49.30
S Containing	26	0.90	0.83	0.69	0.87	3.74	4.00
Cl Containing	27	2.05	3.07	2.24	3.72	7.26	11.86
F Containing	5	1.04	10.14	0.47	11.98	2.33	30.10
Br Containing	8	2.92	2.96	3.56	2.78	9.01	8.40
I Containing	4	1.21	0.678	1.39	0.92	2.78	1.90
P Containing	1	-	-	0	17.20	0	17.20
Multifunctional	78	2.40	9.68	2.25	8.02	8.16	44.80
Total	799	2.53	7.76	1.71	4.10	9.82	56.60

[a] Estimation of $\Delta_f H^o_{gas}$ is performed using the NIST® Structures and Properties software program.

Table 4.26 Comparison of the distribution of data-points based on the range of absolute deviations from the developed GC model and from the Benson's method (1993)

Class of components	Data-points	(0-1) kJ/mol		(1-4) kJ/mol		(4-6) kJ/mol		> 6 kJ/mol	
		This work	Benson's method [9]	This work	Benson's method [9]	This work	Benson's method [9]	This work	Benson's method [9]
Hydrocarbons	259	121	107	116	98	9	19	12	34
O Containing	256	122	82	102	99	24	32	8	43
N Containing	135	64	53	55	42	16	7	-	33
S Containing	26	18	16	8	10	-	-	-	-
Cl Containing	27	9	6	12	11	4	4	2	6
F Containing	5	4	-	2	-	-	-	-	5
Br Containing	8	2	2	3	4	3	2	-	-
I Containing	4	2	3	2	1	-	-	-	-
P Containing	1	1	-	-	-	-	-	-	1
Multifunctional	78	36	22	22	11	11	11	9	35
Total	799	379	291	322	276	67	75	31	157

It can be seen from Table 4.25 and Table 4.26 that the developed GC model for $\Delta_f H^o_{gas}$ provides better prediction performance for the class of organic components containing C, H, O, N, S, Cl, F, and P atoms. Additional experimental data of organic components containing F, Br, I, and P atoms would allow making a fair comparison between the developed GC model and the Benson's method (1993). It is to be noted that, some of the data-points in the data-set used for the parameter regression in this work may not be present in the data-set that was used to fit the Benson's method (1993). Since the information about the data-set that was used to fit the Benson's method (1993) is not available, unfortunately, it cannot be identified whether a particular value of $\Delta_f H^o_{gas}$ calculated using Benson's method (1993) is an estimated value or is a pure prediction.

It is also to be noted that for the purpose of comparing the developed GC model with the Benson's method (1993), the NIST® Structures and Properties software program is used since this program is publicly available. The ASTM CHETAH® 9.0 program which is a commercial software developed by University of South Alabama, Mobile Alabama 36688-0002 and ASTM International, provides estimates of $\Delta_f H^o_{gas}$ based on the revised and latest model parameters of Benson's method. A limited trial version of the ASTM CHETAH® 9.0 program is used to estimate the $\Delta_f H^o_{gas}$ of 30 organic components from the data-set and compare with the absolute deviations with those obtained from the developed GC model (see Table 4.27). The absolute deviations for these 30 organic components obtained from the NIST® Structures and Properties software program are also given in Table 4.27.

The prediction performance of the developed GC model is also compared (in terms of AAE) with currently used quantum mechanical methods. The semi-empirical methods such as the PDDG/PM3 (Repasky et al., 2002) and the PDDG/MNDO (Repasky et al., 2002) which employs Pairwise Distance Directed Gaussian (PDDG) modification, have AAE of about 8-13 kJ/mol (Van Speybroeck et al., 2010) for a data-set of 622 organic components containing C, H, N, and O atoms. The semi-empirical QM method, PM7 in MOPAC2012, yields an AAE of 19 kJ/mol (Van Speybroeck et al., 2010) in $\Delta_f H^o_{gas}$ for C, H, N, O components. The post Hartree-Fock methods (Slater, 1951), on the same data-set of 622 organic components, have AAE's in the order 13-17 kJ/mol (Van Speybroeck et al., 2010). For more accurate estimation of $\Delta_f H^o_{gas}$ (AAE's less than 8 kJ/mol), composite methods such as, G2 (Curtiss et al., 1991), G3 (Curtiss et al., 1998), W1 (Martin and Oliveira., 1993), W2 (Boese et al., 2004), or other variants are often employed (Van Speybroeck et al., 2010). In these methods, the error in the calculation of $\Delta_f H^o_{gas}$ scales with the size of the molecule, so what is within chemical accuracy for very small molecules is far from within that accuracy for medium sized molecules of interest to the chemist and technologist. This implies that apart from the fact that, G2 (Curtiss et al., 1991), G3 (Curtiss et al., 1998), and similar models are black box models, the accuracy of calculation of $\Delta_f H^o_{gas}$ decreases proportionally to the size of

the molecule. Based on AE's of these quantum mechanical methods, it can be seen that the developed model for $\Delta_f H^o_{gas}$ exhibits a much better accuracy with AAE of 1.75 kJ/mol.

Table 4.27 Comparison of prediction performance between the developed GC model and the Benson's method using the ASTM CHETAH® software and the NIST® software for selected organic components

component name	Absolute deviation from developed GC model, kJ/mol	Absolute deviation from ASTM CHETAH software, kJ/mol	Absolute deviation from NIST, kJ/mol
N,N-Dimethylaniline	5.67	8.03	7.10
2-Methylpiperidine	5.69	4.45	0.45
2-Ethyl-1-butanol	5.79	9.62	4.00
1,1'-Biphenyl	5.81	3.01	3.42
2-Propanol, 1-methoxy-	5.97	0.35	1.10
DIPROPYLENE GLYCOL	5.97	24.83	26.20
1,2-Ethanediol	6.04	4.01	3.40
2-Oxepanone	6.26	6.27	6.32
HEXANE, 2,4,4-TRIMETHYL-	6.28	17.93	4.19
Cycloheptane	6.29	1.90	2.40
Pentane, 2,2,4-trimethyl-	6.39	12.83	5.70
2,2,4,4-Tetramethylpentane	6.39	27.33	11.10
1-BUTENE, 2,3,3-TRIMETHYL-	6.44	6.63	8.60
Thiophene, 2,5-dihydro-, 1,1-dioxide	6.47	12.84	0.40
Acetic acid, 1-methylethyl ester	6.47	1.51	3.20
Ethanamine, N,N-diethyl-	6.79	7.53	8.35
Ethanol, 2-amino-	6.81	6.98	6.70
Cyclooctanone	6.95	9.07	5.40
Benzene, 2,4-dichloro-1-methyl-	7.16	13.85	11.86
Ethane, 1,1,2-trichloro-	7.26	0.27	2.00
Benzamide	7.38	8.94	8.94
2-Propenoic acid, methyl ester	7.43	24.01	23.00
1,3-Dimethylcyclohexane	7.55	2.19	1.77
CYCLOPENTANE, 1,1-DIMETHYL-	7.98	0.01	5.28
CYCLOPENTENE, 3-METHYL-	8.14	0.25	0.10
1,3-Cyclohexadiene	8.39	4.18	4.42
Acetic acid, 1,1-dimethylethyl ester	8.85	6.39	6.30
Butane, 2-bromo-	9.01	0.11	0.80
dimethyl-ether	9.73	3.36	2.60
Cyclononane	9.82	1.54	1.50

4.3.1.5 Predictive power of developed GC model

To test the predictive capability of the developed GC model for $\Delta_f H^o_{gas}$, the experimental and predicted values of $\Delta_f H^o_{gas}$ of various classes of organic components are plotted as a function of their carbon number (see Figure 4.7). It can be observed from Figure 4.7 that the predicted values of $\Delta_f H^o_{gas}$ are in good agreement with the experimentally measured values of $\Delta_f H^o_{gas}$ for various classes of organic components and that the developed GC model for $\Delta_f H^o_{gas}$ can be used for safe extrapolation with confidence. Further, to test the reliability of the developed GC model for $\Delta_f H^o_{gas}$, the predictions of $\Delta_f H^o_{gas}$ for 20 organic components are analysed (see Table 4.28) that are not present in the data-set used for the parameter regression. These 20 data-points should not be considered as a validation set (which by definition is to be formed by randomly selecting a fraction of the total data-set), but as additional data-points available for testing the reliability of the predictions from the developed GC model. For these 20 organic components, the developed GC model yields AAE value of 1.99 kJ/mol which is within the required accuracy of ± 2 kJ/mol.

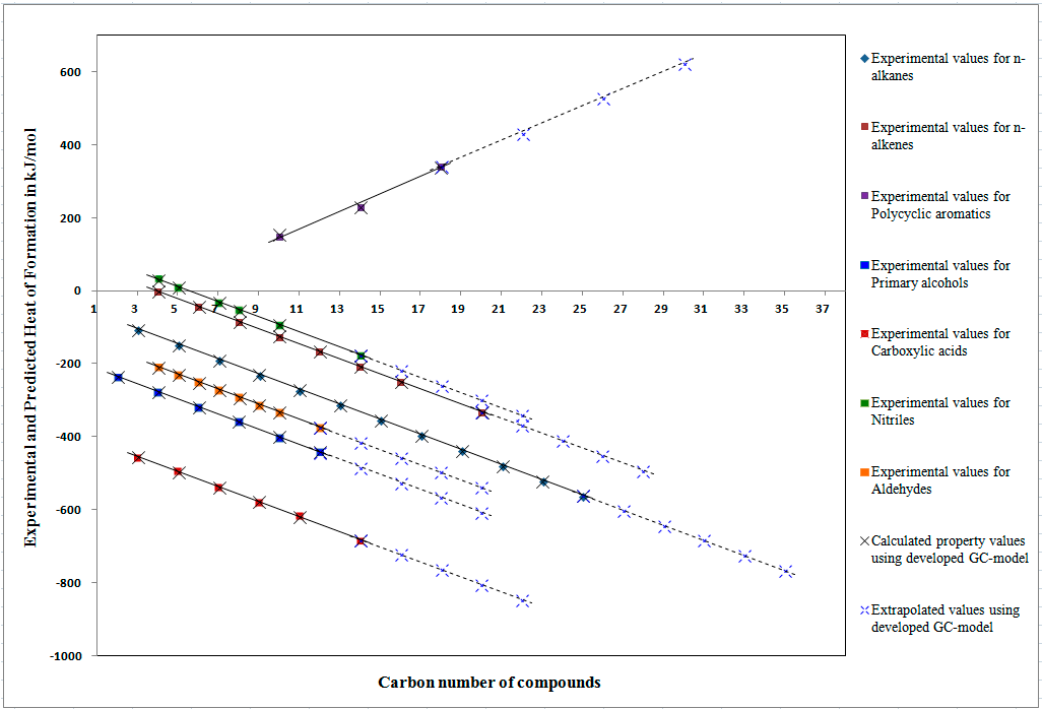


Figure 4.7 Plot of experimental and predicted values of enthalpy of formation versus carbon number of pure components

Table 4.28 Prediction performance for 20 organic compounds not included in the parameter regression step

Component name	Experimental value in kJ/mol	Predicted value from developed GC method in kJ/mol	Absolute deviation from developed GC method in kJ/mol
4-Phenyl-2-butanone	-123.3	-126.28	2.98
1,2,3-Trichlorobenzene	3.8	6.25	2.47
2,6-Dichloropyridine	71.4	71.00	0.40
4,7-Dichloroquinoline	139.4	138.63	0.77
Methoxy 2-chlorobenzene	-99.1	-97.76	1.34
2,3-Dichloroanisole	-118.0	-120.48	2.48
3-Chlorophenylacetic acid	-336.0	-335.29	0.71
Propanedioic acid, diethyl ester	-795.4	-798.52	3.12
Glycine	-390.5	-388.73	1.77
1-Octene, 7-methyl-	-107.7	-111.20	3.50
2-METHYL-1-TRIDECANOL	-487.0	-487.32	0.32
4,4'-Dichlorobiphenyl	120.5	122.63	2.13
2,3-Dichloroquinoxaline	202.9	198.93	3.97
1,1,1-Trifluoro-2,4-pentanedione	-1000.0	-996.01	3.99
Propanenitrile, 2-hydroxy-	-63.9	-63.90	0.00
1-Decylnaphthalene	-66.3	-68.01	1.71
2-Hexanone, 4-methyl-	-305.6	-306.45	0.85
3-Methyl-2-pentanone	-284.1	-281.35	2.75
1,1,1-trichlorobutane	-183.42	-183.68	0.26
2-propanethiol, 2-methyl-	-126.9	-122.56	4.34
AAE			1.99

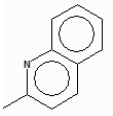
While the developed GC model provides an accurate estimation of $\Delta_f H^o_{gas}$ for a broad range of organic components, it should be noted that the developed GC model is not applicable for some molecular-group components in which a group represents the entire structure of the molecule. A list of molecular-group components which is well documented by Kang et al. (2002) is given in Appendix D for which the developed GC model is not applicable. In addition, the developed GC model is not applicable for the class of gaseous molecules in which the groups are actually atoms, and for the class of molecules with charged atoms.

4.3.2 Application of improved GC model of $\Delta_f H^o_{gas}$

The use of developed GC model to estimate $\Delta_f H^o_{gas}$ of organic components is illustrated by considering the prediction of $\Delta_f H^o_{gas}$ for the component: 2-Methylquinoline (CAS No. 91-63-4). The experimental value for $\Delta_f H^o_{gas}$ for 2-Methylquinoline is 159.1 kJ/mol. Table 4.29

provides the information of MG groups used to represent the molecular structure of 2-Methylquinoline, their occurrences, and contributions of each group. Note that for this component, a new third-order group (aC-aN-aC-$(CH3)_m$ or $(OH)_m$ or $(Cl)_m$ in 0, 1) is defined. Using this information and the property model for $\Delta_f H^o_{gas}$, we estimate $\Delta_f H^o_{gas}$ of 2-Methylquinoline as 159.3 kJ/mol and absolute deviation as 0.2 kJ/mol. Without the inclusion of new third-order group, the estimated value of $\Delta_f H^o_{gas}$ for 2-Methylquinoline was 168.8 kJ/mol and the absolute deviation was 9.7 kJ/mol. It can be seen that the inclusion of new third-order group in the GC model has resulted in very good improvement in the estimation of $\Delta_f H^o_{gas}$

Table 4.29 Prediction of $\Delta_f H^o_{gas}$ for 2-Methylquinoline

2-Methylquinoline	Molecular structure	
		
First-order groups	Occurences	Group-contribution
aCH	6	6.7368
aC fused with aromatic ring	2	19.1902
aN	1	63.4219
aC-CH_3	1	-28.7984
No second-order groups are involved		
Third-order groups	Occurences	Group-contribution
AROM.FUSED[2]	1	10.5077
aC-aN-aC-$(CH3)_m$ or $(OH)_m$ or $(Cl)_m$ in 0, 1	1	-6.8457

$\Delta_f H^o_{gas} = 42.2361\ (H_{fo}) + \sum_i N_i C_i + \sum_j M_j D_j + \sum_k E_k O_k = 159.3$ kJ/mol [a]

The experimentally measured value is 159.1 kJ/mol. Hence, the absolute deviation is 0.2 kJ/mol.

[a] Joback and Reid method [5] = 199.2 kJ/mol; Benson's method [9] = 182.0 kJ/mol; Domalski and Hearing method [10] = 170.7 kJ/mol.

To estimate the uncertainty of the predicted $\Delta_f H^o_{gas}$, we use the linear approximation of nonlinear least squares theorem (see section 3.2.4 of chapter 3) which involves the estimation of the covariance $COV(\boldsymbol{P}^*)$ of the involved groups and universal constant H_{f0}, and local sensitivity $J(\boldsymbol{P}^*)$ of the property model. The covariance of the listed groups given in Table

4.30 was extracted from the whole covariance matrix of all the groups involved in the analysis of model for $\Delta_f H^o_{gas}$. In Table 4.30, only lower triangular elements are shown since the upper triangular matrix elements are identical to the lower ones. Table 4.31 lists the local sensitivity of $\Delta_f H^o_{gas}$ model with respect to the model parameters (for contributions listed in Table 4.29 and the universal constant, H_{f0}). It is to be noted that the property model for $\Delta_f H^o_{gas}$ is a linear one and hence, the local sensitivity values, $J(\boldsymbol{P}^*)$, are equal to the occurrences of the groups involved in 2-Methylquinoline.

Table 4.30 Covariance matrix $COV(P^*)$ for the groups listed in Table 4.29

	H_{fo}	aCH	aC	aN	aC- CH_3	AROM.FUSED[2]	aC-aN-aC-$(CH3)_m$ or $(OH)_m$ or $(Cl)_m$ in 0, 1
H_{fo}	10.70						
aCH	-1.78	0.33					
aC	2.05	-0.42	17.78				
aN	-1.80	0.34	-1.41	1.44			
aC- CH_3	-1.77	0.20	0.31	-0.16	0.93		
AROM.FUSED[2]	-0.27	0.00	-17.98	0.94	-0.24	19.95	
aC-aN-aC-$(CH3)_m$ or $(OH)_m$ or $(Cl)_m$ in	-0.25	0.13	-16.58	0.37	-0.65	16.35	18.45

Table 4.31 Local sensitivity ($J(P^*)$ of $\Delta_f H^o_{gas}$ model with respect to model parameters

$\delta f/\delta H_{fo}$	$\delta f/\delta$aCH	$\delta f/\delta$aC	$\delta f/\delta$aN	$\delta f/\delta$aC- CH_3	$\delta f/\delta$ AROM.FUSED[2]	$\delta f/$ δaC-aN-aC-CH3 or OH or Cl
1	6	2	1	1	1	1

To calculate the confidence intervals, say the 95% confidence intervals of the predicted $\Delta_f H^o_{gas}$ value, the covariance matrix $COV(\boldsymbol{P}^*)$ given in Table 4.30 and the local sensitivity $J(\boldsymbol{P}^*)$given in Table 4.31 are substituted in Eq. (3.13). For 95% confidence interval calculation, the t-distribution value corresponding to 0.05*/2* percentile (i.e. α_t /2 percentile) and with 577 degrees of freedom (obtained by deducting 284 estimated parameters from 861 data-points) is 1.96. The calculated 95% confidence intervals of estimated $\Delta_f H^o_{gas}$ value is,

$$X^{pred}_{1-\alpha_t} = X^{pred} \pm \underbrace{\sqrt{\text{diag}\left(J\left(\boldsymbol{P}^*\right)COV\left(\boldsymbol{P}^*\right)J\left(\boldsymbol{P}^*\right)^T\right)}}_{1.27} \cdot \underbrace{t\left(\nu,\alpha_t/2\right)}_{1.96} = (159.3 \pm 2.5)\ \text{kJ/mol}$$

It can be observed that the experimental value (159.1 kJ/mol) falls in between the predicted confidence intervals indicating reliability of the developed model for $\Delta_f H^o_{gas}$.

4.3.3 Application of molecular structure similarity criteria to modeling of other pure component properties

Using the 'molecular structure similarity criteria' based approach, it has been possible to improve the performance of models for a wide range of properties of pure components. The $\Delta_{fus}H$ of pure components is a key thermodynamic property in predicting solubility in aqueous/organic solvents for which both the measured data as well as GC based models for organic components have been reported (Marrero and Gani, 2001; Chickos and Acree Jr., 2009; Chickos et al., 1998). The GC model for $\Delta_{fus}H$ developed in this work (see section 4.1 of this chapter) has SD of 4.92 kJ/mol and AAD of 2.58 kJ/mol. The accurate prediction of T_c of pure components is also important since it is used in many engineering calculation involving use of equation-of-state based models. The GC model for T_c developed in this work (see section 4.1 of this chapter) has SD of 10.91 K and AAD of 7.28 K. To improve the prediction performance of GC models for $\Delta_{fus}H$ and T_c, the newly defined third-order groups (see Table 4.21) were applied to the components in the data-set of $\Delta_{fus}H$ and T_c and the regression analysis is performed. Table 4.32 provides the comparison between the model performance statistics before and after the application of molecular similarity structural criteria approach. The total list of first-order, second-order and third-order groups and their contributions C_i, D_j, and E_k for the improved GC models for $\Delta_{fus}H$ and T_c are given in Appendix C (see Tables C1-C3). These improvements obtained therefore add credits to the validity of newly defined third-order groups developed above while modeling of $\Delta_f H^o_{gas}$ property using 'molecular structural similarity criteria' approach.

Table 4.32 Comparison of performance of GC models for $\Delta_{fus}H$ and T_c

Property	Data-points	Model performance statistics based on the previous GC model parameter values (see Table A1 of Appendix A)					Model performance statistics after employing similarity structural criteria approach				
		R^2	SD	AAE	ARE[a]	AE_{max}	R^2	SD	AAE	ARE[a]	AE_{max}
T_c in [K]	758	0.99	10.91	7.82	1.26	44.34	0.99	9.06	6.35	1.02	38.13
$\Delta_{fus}H$ [kJ/mol]	719	0.84	4.92	2.58	--	69.09	0.93	3.26	2.0	--	31.4

[a] For $\Delta_{fus}H$, ARD is not reported since this property contain both positive and negative values.

The reduced Chi-squared statistics are calculated (see Table 4.33) and are used as an additional criterion to compare the performance of improved GC models using the 'molecular structural similarity criteria' approach with the previous GC models developed in this work. The expression for reduced Chi-square χ^2_{red} is given by (Taylor, 1997):

$$\chi^2_{red} = \frac{N}{N-P-1}\left(\frac{\sigma_d}{\Delta}\right)^2 \tag{4.4}$$

Where N is number of data-points; P = number of parameters estimate; σ_d = standard deviation; and Δ = error of measurement.

Table 4.33 Chi-squared reduced statistics for the developed and previous version of GC models for $\Delta_f H^o_{gas}$, $\Delta_{fus}H$ and T_c

Property	χ^2_{red} for the previous GC models					χ^2_{red} for the improved GC models				
	N	P	σ_d	Δ	χ^2_{red}	N	P	σ_d	Δ	χ^2_{red}
$\Delta_f H^o_{gas}$ in [kJ/mol]	882	233	6.6	2.0	14.85	861	286	2.61	2.0	2.55
T_c in [K]	858	251	10.91	5.0	6.74	873	287	9.06	5.0	4.89
$\Delta_{fus}H$ in [kJ/mol]	764	248	4.92	2.0	8.97	719	245	3.26	2.0	4.03

Note that the assumed values of error of measurement, Δ, are based on the typical values of uncertainties of experimentally measured property values. The comparison in Table 4.32 shows that the reduced Chi square values for improved new models are much better.

4.3.4 Selection of minimum data-set for the parameter regression

The methodology discussed in section 3.2.1 of chapter 3 is applied to find the minimum data-set for the parameter regression in the property modeling of $\Delta_f H^o_{gas}$. The total data-set of $\Delta_f H^o_{gas}$ contains measured values for 861 (that is, m) pure components. The total number of unknown GC model parameters (first-order, second-order, and third-order MG groups) is 285 (that is, n). The occurrence matrix, A, is therefore of size $m \times n$. The ρ(A) for this case is 285 and ρ(A|**b**) is 286. By following the procedure as suggested in section 3.3.1 of chapter 3, it is found that the minimum data-set consisting of 806 data-points results in an augmented matrix rank of 285. The minimum data-set of 806 data-points is obtained by analyzing the trends of the $\Delta_f H^o_{gas}$ of a particular class of components and selecting the data-points (from the interpolation range) that are to be retained in the minimum data-set. The parameter regression is then performed ensuring that Eq. (3.6) is satisfied. Using this minimum data-set, all of the 285 unknown GC model parameters are estimated successfully thus achieving same application range of the model as it is achieved when regression is performed using the total

data-set. The remaining 55 data-points consisting of components from various classes such as *n*-alkanes, *n*-alkenes, primary alcohols etc. are used for the validation purpose. Table 4.34 provides the summary of performance statistics for the total data-set, for the minimum data-set and for additional data-points used for the model validation. It can be observed from Table 4.33 that the selected minimum data-set is able to provide similar performance statistics as those obtained by regression of the total data-set. Also, for the extra 55 data-points selected for the validation, the predictions are accurate and reliable. These statistics helps to assess the validity of the method for finding the minimum data-set as well as to assess the reliability and prediction ability of the improved GC model for $\Delta_f H^o_{gas}$. In this work, the proposed method for selecting minimum data-set is tested only for the modeling of enthalpy of formation. However, it is to be noted that this method, as a future work, needs to be further applied and tested in the property modeling of other pure component properties. Further, the effect of measurement error on the selection of the minimum data-set for regression should be investigated.

Table 4.34 Performance statistics for the total data-set, for the minimum data-set and for additional data-points used for the model validation.

Data-set	R^2	SD in kJ/mol	AAE in kJ/mol
Total data-set (861 data-points)	0.99	2.61	1.75
Minimum data-set (806 data-points)	0.99	2.68	1.82
55 data-points for the validation [a]	-	1.19	0.94

[a] R^2 is not given since the data-points are considered only for the evaluation purpose

4.3.5 Summary

A molecular structure similarity criteria based approach is applied to improve the performance of GC models for enthalpy of formation, enthalpy of fusion, and critical temperature by making efficient use of knowledge of molecular structural information and properties of pure components. The methodology helps to achieve the required level of prediction accuracy of pure component property models. For all the properties listed above, significant improvements in the performance of their property models are achieved. Compared to the currently used GC models and quantum mechanical methods, the developed GC model for $\Delta_f H^o_{gas}$ performs much better and, importantly, within a *chemical accuracy*. With respect to the quantity of data to be used for the parameter regression, in general, it is advisable to use all of the available experimental data-points for regressing GC model parameters. However, if it is preferred to retain some of the data-points for the model validation purpose then the developed method for selecting minimum data-set for the

parameter regression can be used. This can be found from identifiability of linear systems which relies on the rank analysis of the augmented matrix. The selection of minimum data-set for the parameter regression allows the model developer to achieve comparable correlation statistics as well as to validate the reliability of pure predictions.

Chapter 5. Results: Software Development

5.1 Introduction to ProPred

The GC^+ method based property models are simple, fast, and do not require substantial computational resources to calculate and obtain the needed pure component property values. An automation of GC^+ method based property models makes the property estimation even faster and convenient for the property user. Therefore, the developed GC^+ method based property models for the estimation of thermo-physical, transport-related, and environmental-related properties of pure components are implemented in the ProPred software. ProPred is a property estimation toolbox integrated within the Integrated Computer Aided System software (ICAS® 16, User manual) developed at CAPEC, DTU. The ProPred software is build based on the Visual C++ 6.0 environment. A graphical user interface is available within the ProPred so that the user can provide the molecular structure information of pure components either by using available drawing tools within the ProPred or by providing SMILES information of molecules. Using the molecular structural of pure component as input information, ProPred then predicts pure component properties based on the incorporated GC methods developed by Marrero and Gani (2001), Constantinou and Gani (1994), and Joback and Reid (1987). ProPred also contains a feature that allows the user to automatically employ GC^+ approach to estimate properties of pure components whose molecular structure is not completely described by the available Marrero and Gani groups. The CAPEC database (with more than 23000 molecules) containing available experimental data of wide range of pure components is also available in the ProPred.

The main window of ProPred software is divided into two sub-windows. In the right sub-window which is the drawing area, the user needs to provide the input information as the molecular structure of the pure component whose properties are to be estimated (see Figure 5.1). This can be done by either using the drawing tools or by importing SMILES of the pure component. The left sub-window which is the property area, displays the estimated values of properties of organic as well as polymer molecules using several incorporated GC methods. There are several useful features available in the menu bar of ProPred as shown in Figure 5.2.

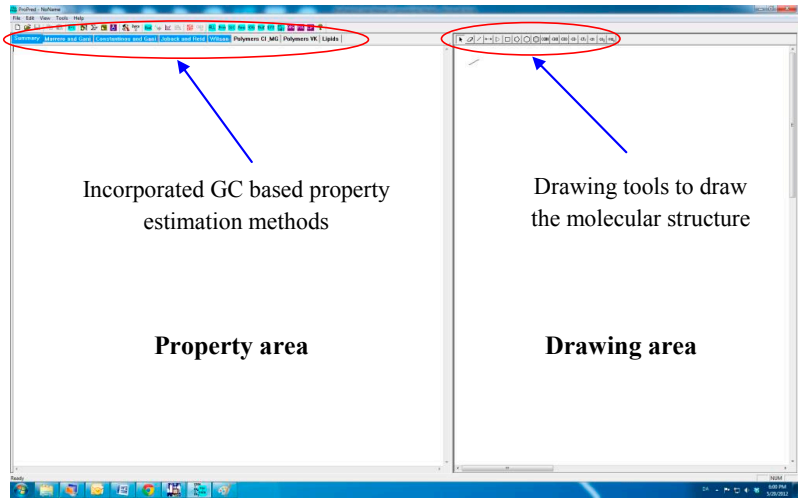


Figure 5.1 Starting window in ProPred

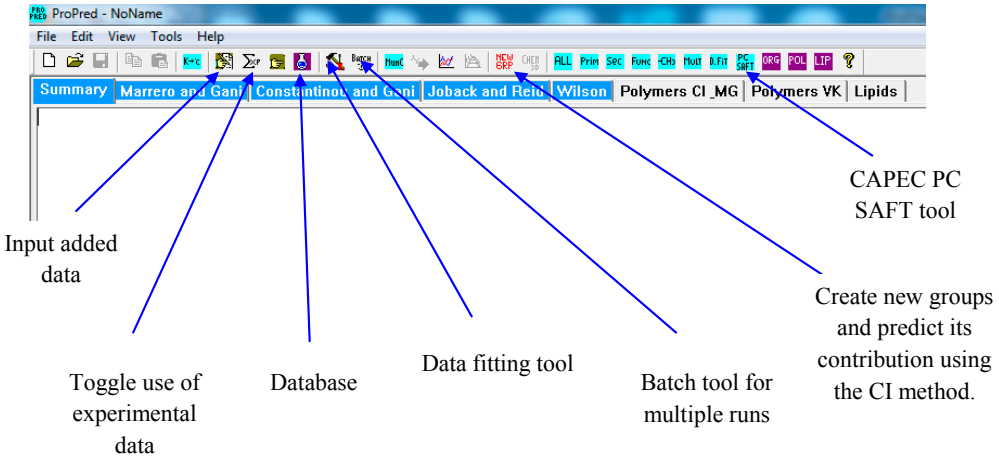


Figure 5.2 ProPred menu bar

5.2 Inclusion of Developed GC$^+$ Based Property Models for Thermo-Physical and Transport-Related Properties in the ProPred

The developed GC$^+$ based property models for the estimation of thermo-physical and transport-related properties of pure components are implemented in the 'Marrero and Gani' page of ProPred. Figure 5.3 shows a screen shot of 'Marrero and Gani' page displaying the estimated property values for the pure component, Butanedoic acid, dipropyl ester.

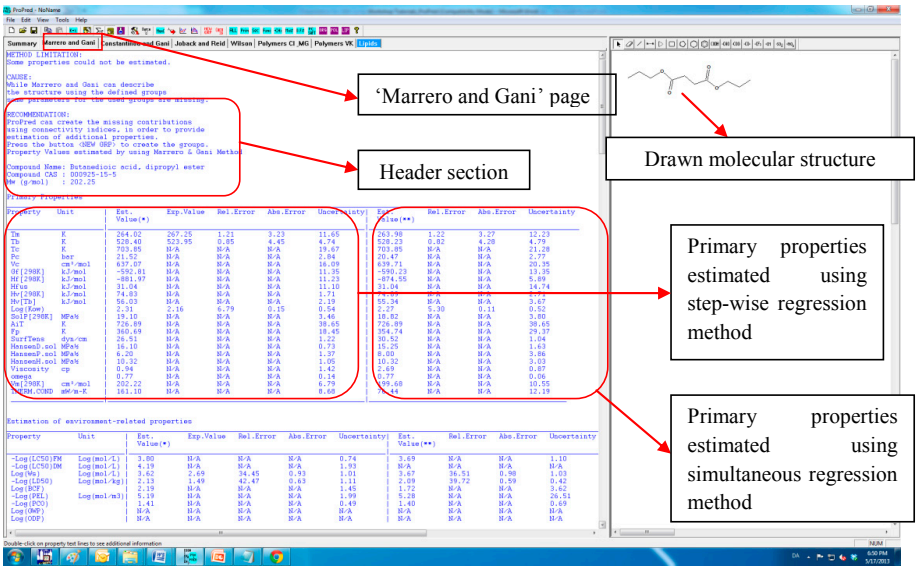


Figure 5.3 A screenshot of 'Marrero and Gani' page of ProPred

The first section of the 'Marrero and Gani' page is a header with the chemical name, CAS number and molecular weight of the drawn molecule. Both the name and the CAS number are either user-supplied or retrieved from an internal database. The next section is the 'Primary Properties' section. Primary properties are those which only depend on the molecular structure. For each property, the following data is displayed.

- Short identifier for each pure component property (for example, *Tb*, *Tc*, *Pc*, etc.)
- Units of the property value (for example, K, kJ/mol, cc/mol, etc.)
- Estimated property value using Marrero and Gani GC / GC^{+} method
- Experimental value when available in the internal database of ProPred
- Absolute error (when the experimental value is available) calculated using the formula: | (experimental value – predicted property value) |.
- Relative error (when the experimental value is available) calculated using the formula:100 * | (experimental value – predicted property value) | / experimental value.
- The estimates of uncertainties of predicted property values as 95% confidence intervals quantified using the maximum likelihood approach (covariance matrix based approach as discussed in section 3.2.4 of chapter 3).

Further, the property estimation based on model parameters determined using the step-wise regression method and using the simultaneous regression method has been included. The last section of the 'Marrero and Gani' page is the group assignments section which displays the group description of the drawn molecular structure (see Figure 5.4).

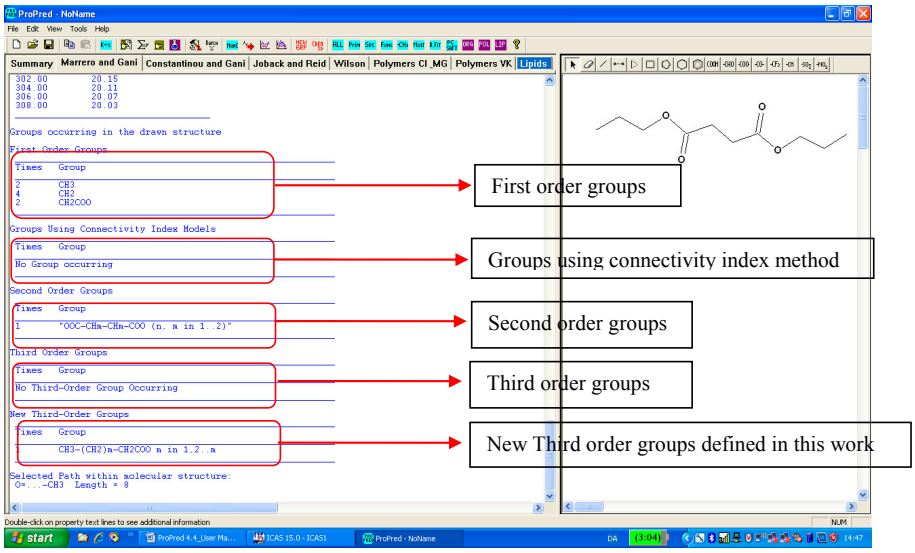


Figure 5.4 A screenshot of group assignments section in the 'Marrero and Gani' page

5.2.1 Application of GC$^+$ approach

Figure 5.6 illustrates a situation in which a PH_2 group is attached to the molecular structure of Butanedoic acid, dipropyl ester. Note that the PH_2 group is not available in the existing list of first-order Marrero and Gani groups and hence the property estimation using that GC method is not possible for this component. After attaching the PH_2 group, a message has been displayed in the property display area of ProPred stating that "METHOD LIMITATION: Unable to estimate any property by using the current set of groups" (see Figure 5.5). To overcome this limitation, the user needs to employ the GC$^+$ approach. To employ the GC$^+$ approach, the toolbar button NEW GRP is to be clicked after which the display appear as shown in Figure 5.6.

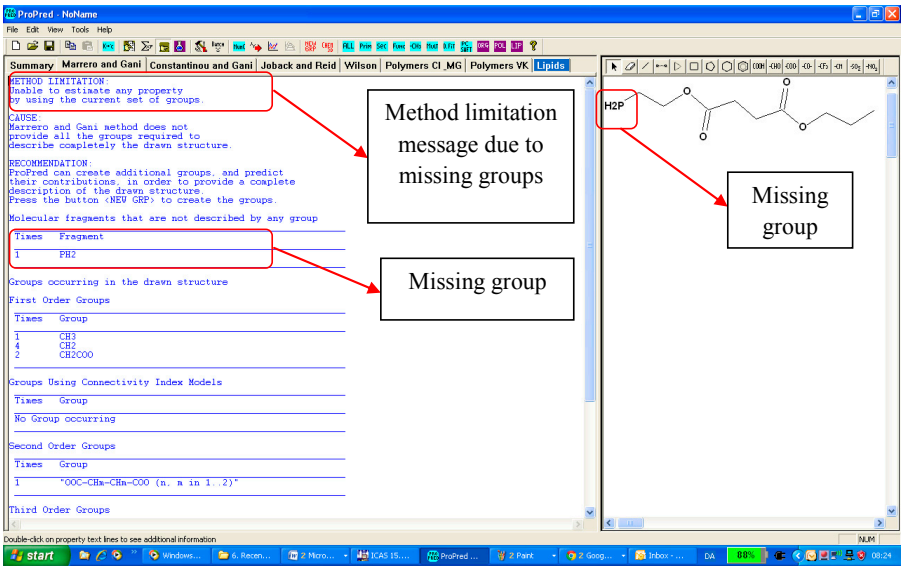


Figure 5.5 A screenshot of molecular structure involving a missing group (PH_2)

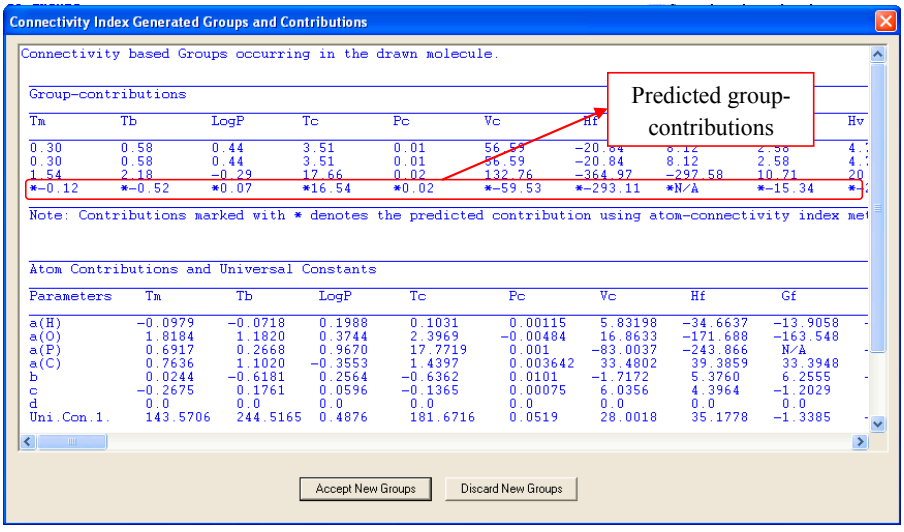


Figure 5.6 A screenshot of predicted group-contributions values for the missing groups

The contribution values of the missing group (PH_2) are shown in the dialog (these values are marked with stars). The atom contributions, the zeroth-order and first-order connectivity indices, and CI model parameters *b* and *c* for different pure component property models as developed in this work have been included in the ProPred for the purpose of creating missing groups and predicting their missing group contributions. After pressing the button 'Accept

New Groups', the missing groups and their contributions are included in the property estimation using the MG method (see Figure 5.7).

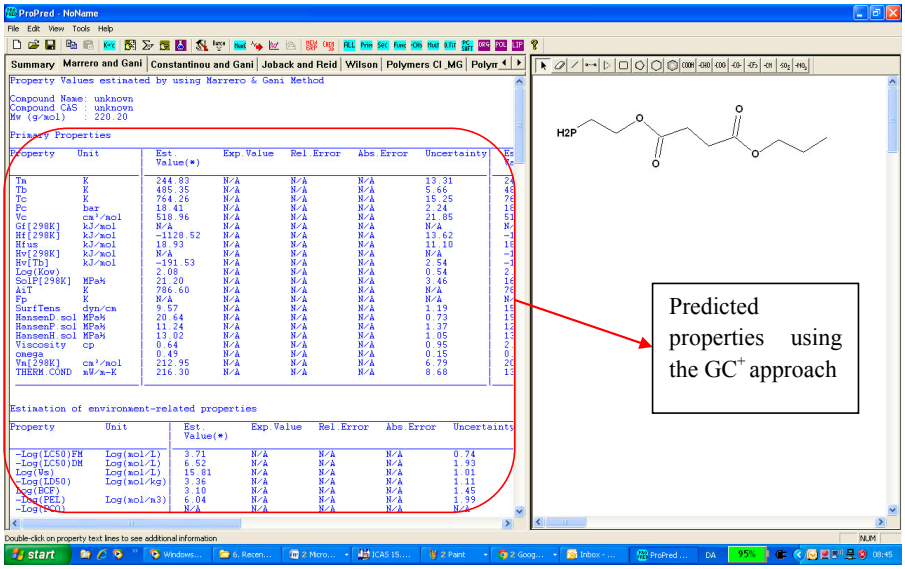


Figure 5.7 A screenshot of predicted properties using the GC$^+$ approach

5.3 Inclusion of Developed GC$^+$ Based Property Models for Environmental-Related Properties in the ProPred

In the ProPred, the developed GC$^+$ based property models for the environmental-related properties of organic chemicals have been included in the 'Primary Properties' section of the 'Marrero and Gani' page. These property models provide necessary property estimates together with uncertainties of estimated values needed in performing the sustainability analysis using the WAR® algorithm and the USEtox® model. The first 10 properties listed in the table 'Estimation of environmental-related properties' (see Figure 5.8) are necessary to perform the sustainability analysis using the WAR® algorithm. These properties include, fathead minnow 96-hr LC_{50} (-Log(LC50)FM), *daphnia magna* 48-hr LC50 (-Log(LC50)DM), oral rat LD_{50}, aqueous solubility ($LogW_s$), bio-concentration factor (Log(BCF)), permissible exposure limit (-Log(PEL)), photochemical oxidation potential (-Log(PCO)), global warming potential (Log(GWP)), ozone depletion potential (Log(ODP)), and acidification potential (Log(AP)).

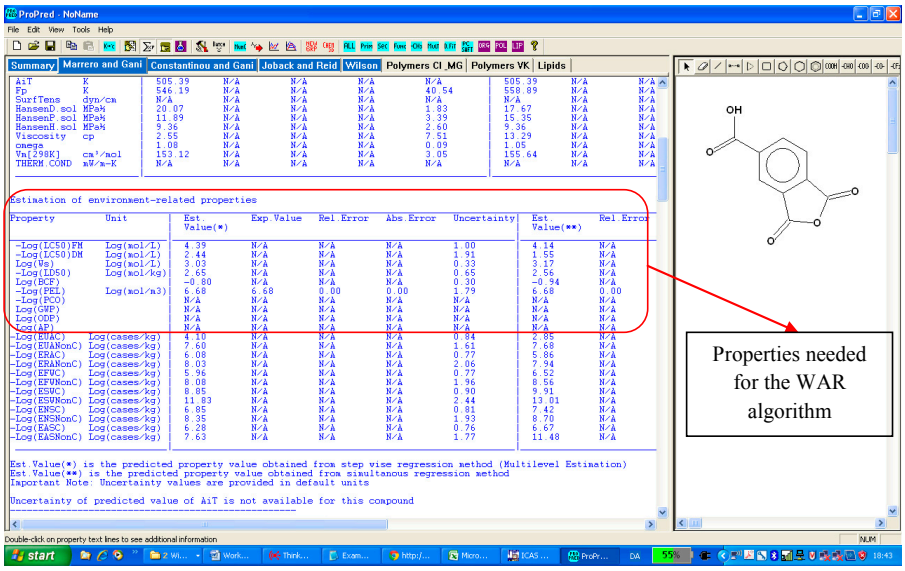


Figure 5.8 A screenshot of estimation of environmental-related properties needed for the WAR algorithm

The next 12 properties listed in the table “Estimation of environmental-related properties” (see Figure 5.9) are necessary in performing the sustainability analysis using the USEtox® model.

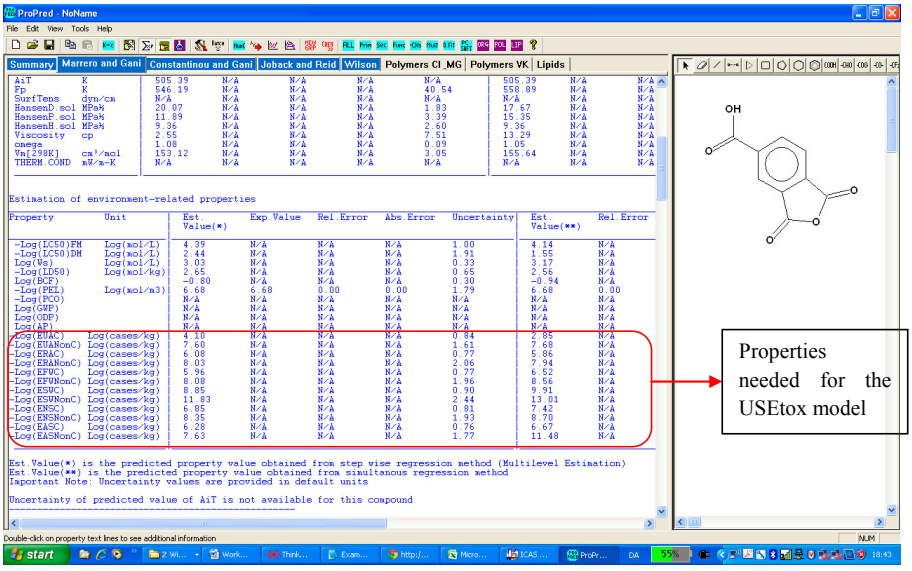


Figure 5.9 A screenshot of estimation of environmental-related properties needed for the USEtox model

These properties include, emission to urban air (carcinogenic, -Log(EUAC) and non-carcinogenic, -Log(EUANonC)), emission to continental rural air (carcinogenic, -Log(ERAC) and non-carcinogenic, -Log(ERANonC)), emission to continental fresh water (carcinogenic, -Log(EFWC) and non-carcinogenic, -Log(EFWNonC)), emission to continental sea water (carcinogenic, -Log(ESWC) and non-carcinogenic, -Log(ESWNonC)), emission to continental natural soil (carcinogenic, -Log(ENSC) and non-carcinogenic, -Log(ENSNonC)), emission to continental agricultural soil (carcinogenic, -Log(EASC) and non-carcinogenic, -Log(EASNonC)).

5.4 Inclusion of improved GC property models developed using molecular structure similarity criteria in the ProPred

For enthalpy of formation, enthalpy of fusion, and critical temperature, the improved MG method based property models have been included in the ProPred. The improved property models for these properties involve use of newly defined third-order Marrero and Gani groups (see section 4.3 of chapter 4) in the property estimation. If the drawn molecular structure of a molecule involves any of these newly included third-order Marrero and Gani groups, then ProPred takes into account the contribution of these new third-order groups in the property estimation. Also, the definition of new third-order groups is displayed under the group assignment section of the 'Marrero and Gani' page under the heading 'New Third Order Groups'. For illustration, let us consider the prediction of enthalpy of formation of the pure component, DIETHYLENE GLYCOL METHYL ETHER. Figure 5.10 shows a screen shot of 'Marrero and Gani' page displaying the estimated property values for the pure component, DIETHYLENE GLYCOL METHYL ETHER.

The enthalpy of formation for DIETHYLENE GLYCOL METHYL ETHER can be obtained by viewing at 'Hf[298K]' in the property display area of ProPred (see Figure 5.10). The experimental value of enthalpy of formation for DIETHYLENE GLYCOL METHYL ETHER as retrieved from the ProPred database shows a value of -559 kJ/mol. The Predicted value of enthalpy of formation for DIETHYLENE GLYCOL METHYL ETHER using the improved GC model is -559 kJ/mol and hence the relative error and absolute error is zero. The improvement in the prediction of enthalpy of formation is due to the inclusion of new third-order group in the MG method based property model. Before inclusion of the new third-order group in the property estimation, the estimated value of enthalpy of formation for DIETHYLENE GLYCOL METHYL ETHER is, -539.1 kJ/mol and hence the absolute deviation was 19.85 kJ/mol.

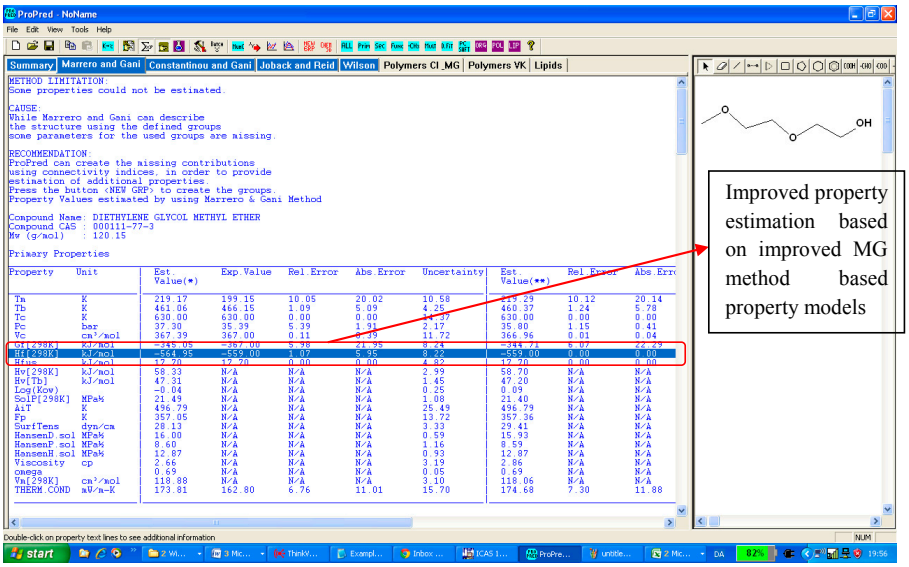


Figure 5.10 A screenshot of estimation of enthalpy of formation using improved GC model

The Marrero and Gani group assignments for DIETHYLENE GLYCOL METHYL ETHER can be viewed at the bottom of 'Marrero and Gani' page of ProPred (see Figure 5.11). It can be seen that for DIETHYLENE GLYCOL METHYL ETHER, a new third-order group, CH3O-(CH2)m-OCH2CH2OH, has been included in the property estimation.

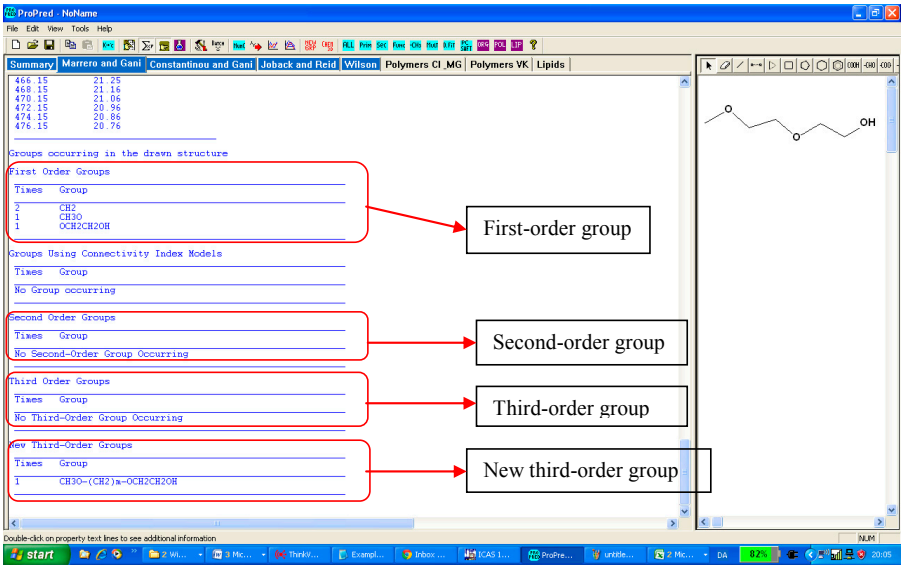


Figure 5.11 A screenshot of new third-order group displayed in ProPred

A complete list of pure components considered in the performance improvement of GC model for $\Delta_f H^o_{gas}$ (see section 4.3 of chapter 4) and the residuals obtained from the regression analysis and from the ProPred software are provided in the supplementary material which can be downloaded from the following link: http://www.capec.kt.dtu.dk/documents/enthalpy_of_formation/Comparison_Between_Residuals_From_Regression_and_From_ProPred.pdf. The experimentally measured values of $\Delta_f H^o_{gas}$ of organic compounds used in the regression are not given due to confidentiality reasons. An erratum mentioning discrepancies observed in the estimated value of enthalpy of formation from the improved GC model for $\Delta_f H^o_{gas}$ and from the ProPred is given in this supplementary material.

5.5 Summary

The following ProPred software development work has been done in this work.

(i) Inclusion of developed GC^+ based property models for the estimation of thermo-physical and transport-related properties of pure components
(ii) Inclusion of developed GC^+ based property models for the estimation of environmental related properties of pure components
(iii) Inclusion of improved GC based property models developed using molecular structure similarity criteria based approach for the estimation of enthalpy of formation, enthalpy of fusion, and critical temperature of pure components. There is a need to develop and implement a method in the ProPred so that the new third-order groups in the drawn molecular structure of pure component can be identified accurately.

These implemented property models allow the user to make fast and reliable property estimation of thermo-physical, transport-related, and environmental-related properties of pure components needed in the chemical product-process design and sustainability analysis. The user needs to provide only the input information as the molecular structure of pure components whose properties are to be estimated.

Chapter 6. Results: Applications in Process Design

In this chapter, effects of uncertainties of pure component property estimates on the process design are discussed through the proposed methodology for performing sensitivity analysis of process design due to uncertainties of property estimates (see section 3.2.6 of chapter 3). To illustrate the application of this methodology, the following three case studies are considered:

(1) An extractive distillation process to separate acetone from methanol using water as an entrainer
(2) Short-path evaporator to separate glycerol from the mixture of acylglycerides
(3) A de-acidification system of a vegetable oil deodorization process to recover free fatty acids (FFA's) and nutritional components (Tocopherols, Tocotrienols etc.)

To perform the sensitivity analysis, the properties of pure components involved in the process model are perturbed and then process design calculations are carried out to quantify the effects of these uncertainties on the design variables. It is to be noted that, while performing sensitivity analysis in these case studies, effects of uncertainties of predicted temperature-dependent properties of pure component such as, liquid vapor pressure, heat of vaporization, liquid density, liquid viscosity, and liquid surface tension are also considered for the sake of completeness. Inclusion of these properties in the analysis helps in understanding the impact of prediction errors of these properties on the quality of process design. To perform the process design calculations for case studies 1 and 3, the PRO/II® simulator is used. To perturb property values of temperature dependent properties, the following procedure is adopted: First, the property values are perturbed to a desired value of uncertainty (say, +1%, +2% etc.) and using the perturbed data, parameter regression is performed to obtain new parameter values of correlations for temperature dependent properties. Then these new parameters are entered into the property library of PRO/II® to achieve the desired perturbation of a property value. To perform the process design calculations and sensitivity analysis for the case study 2, MoT toolbox of ICAS® is used. When prediction of pure component properties using developed GC^+ models is required, the predicted property values together with uncertainty estimates (in terms of 95% confidence intervals) are taken from the ProPred. The uncertainty estimates of predicted temperature dependent properties of pure components estimated using available DIPPR® correlations are obtained from the DIPPR® database. The uncertainty estimates (in terms of average relative error) of predicted primary and temperature-dependent properties of pure components are used together with the results of sensitivity analysis to identify the most important properties from process design point of view and to list and rank them based on their critical importance.

6.1 Case Study 1: Sensitivity Analysis of Design of Extractive Distillation Process due to Uncertainties of Property Estimates

Extractive distillation is used to separate pure components from a mixture with low relative volatility forming an azeotrope (Henley and Seader, 2006). A miscible entrainer/solvent is used to increase the relative volatilities between the key components to be separated. The choice of the best solvent is critical for a feasible, economic and environmental friendly operation. The selection of a suitable solvent is based on various criteria such as, selectivity, thermal stability, low toxicity, easy recovery and cost (Gil et al., 2009). An extractive distillation column can be divided into three sections where the section between the top of the column and the solvent feed stage is defined as the rectifying section. The section between solvent feed stage and the feed stage for the azeotropic mixture is defined as the extractive section and the section between feed stage and the bottom of the column is defined as the stripping section. After the separation task in the extractive distillation column, the entrainer is recovered in a second conventional distillation column and then recycled back to the extractive distillation column.

Step 1. Problem definition

It is required to perform sensitivity analysis of design of an extractive distillation process employed to separate acetone from methanol using water as an entrainer to obtain a product of at least 99.4% of acetone subject to uncertainties in the property estimates.

Step 2. Collect base case data

The base case data is collected from Gil et.al. (2009) and is summarized in Table 6.1 for the extractive distillation column and in Table 6.2 for the entrainer recovery column. The involved single value pure component properties and temperature-dependent properties are collected from the SIMSCI library available in PRO/II®. Note that, the experimentally measured values of pure component properties such as critical temperature for acetone, methanol, and water are available in the SIMSCI library and hence, in this case study, these experimental data have been considered in the process design calculations. For estimation of temperature dependent properties of acetone, methanol, and water, the DIPPR® correlations available in the SIMSCI library are used. The UNIQUAC model is employed to calculate the liquid phase activity coefficients while the vapor phase is assumed to behave as an ideal gas. The pure component properties considered in the sensitivity analysis are: (i) vapor pressure (P_{vap}), (ii) critical temperature (T_c), (iii) heat of vaporization ($\Delta_{vap}H$), (iv) liquid density (ρ), (v) liquid viscosity (μ), and (vi) liquid surface tension (σ). The design variables considered for the sensitivity analysis are: (i) number of stages (N), (ii) reflux ratio (R), (iii) entrainer stage (N_s), (iv) column diameter (D), and (v) reboiler (Q_r) and condenser (Q_c) heat duties. The uncertainties (in terms of average relative error) of predicted P_{vap}, $\Delta_{vap}H$, ρ, and σ of

acetone using DIPPR® correlations as taken from the DIPPR® database are ±0.25%, ± 3.44%, ± 0.38%, and ±1.77% resp. The uncertainty in the experimental value of T_c (508 K) of acetone as noted from the NIST® database is ± 2 K (that is, ±0.4%).

Table 6.1 Base case design data for extractive distillation column (Gil et al., 2009)

Specifications	Values
Feed composition (mole fraction)	Acetone = 0.7775; Methanol = 0.2225
0.2225Feed temperature	293.15 K
Feed pressure	1 atm
Entrainer	Pure water
Entrainer temperature	320.15 K
Entrainer pressure	1 atm
Loss of entrainer	0.4215 kmol/hr
Entrainer to feed ratio	2.0
Number of stages	52
Entrainer stage	22
Feed stage	48
Reflux ratio	5.0
Pressure drop per stage	0.01 atm
Thermodynamic property model	UNIQUAC

Table 6.2 Base case design data for entrainer recovery column (Gil et al., 2009)

Specifications	Values
Number of stages	26
Feed stage	14
Reflux ratio	3.0
Methanol composition at the top	88.2 %
Entrainer composition at the bottom	99.9 %
Pressure drop per stage	0.01 atm

Step 3. Construct process model

Figure 6.1 shows a steady state process model for an extractive distillation process as constructed in PRO/II®. The input data necessary for performing process design as collected in *Step 2* is also shown in Figure 6.1.

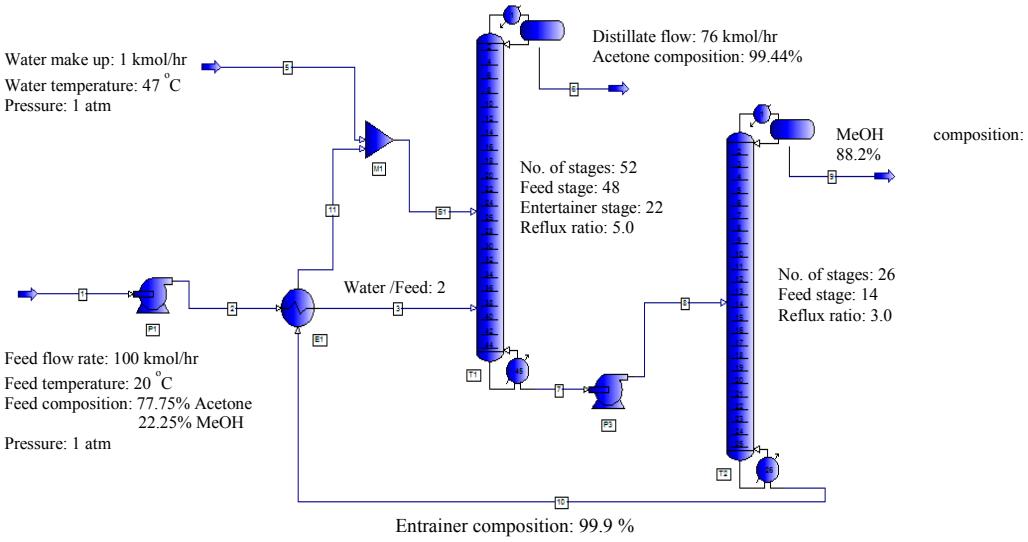


Figure 6.1 Steady state process model for extractive distillation process in PRO/II®

The binary feed mixture containing acetone and methanol is preheated using the heat available in the recycled entrainer stream and is then fed into extractive distillation column at the 48th stage. The water (entrainer) is fed at the 22nd stage. The operating reflux ratio for extractive distillation column is maintained at a base case value of 5.0. The product stream containing acetone of purity of 99.4% is continuously withdrawn from the top of the extractive distillation column. The bottom stream containing methanol and water is fed to entrainer recovery column to separate water from methanol. The separated water is recycled back to the extractive distillation column. The operating reflux ratio for entrainer recovery column is maintained at a base case value of 3.0.

Step 4. Solve, verify, and validate the model

The model is solved in PRO/II® and the obtained product purity of acetone from extractive distillation column and that of methanol from recovery column are compared and validated against the values given by Gil et al. (2009). A good match is obtained. These values are provided in Figure 6.1. This forms the base case design for performing sensitivity analysis of process design due to uncertainties of property estimates.

Step 5. Perform sensitivity analysis

The sensitivity analysis of design of extractive distillation process due to uncertainties of property estimates is performed using the local sensitivity analysis. The results of sensitivity analysis for design variables with respect to uncertainties of predicted properties are shown in Figures 6.2(a)-6.2(g). In these figures, the sensitivity analysis results are reported for only those simulations in which a converged solution satisfying the desired product purity is obtained from PRO/II®. For example, when P_{vap} of acetone is perturbed by a value of more than +1% and less than -0.25%, the converged solution from PRO/II® satisfying the purity of 99.4% of acetone was not achieved and hence these are not reported in Figure 6.2(a). When necessary, chain rule is applied to obtain relationships between the design variables and the properties analysed. For example, to obtain sensitivity of N_s with respect to uncertainties in the predicted P_{vap}, the following chain rule is applied: $[\partial R/\partial P_{vap}]*[\partial N_s/\partial R] = [\partial N_s/\partial P_{vap}]$. In Figures 6.2(a)-6.2(g), the y-axis represents the percent deviation of design variables from their base case values namely, $R0$, $N0$, N_s0, Q_r0, and $D0$. The zero point on the x-axis represents the base case design.

Step 6. Evaluation of results of sensitivity analysis

Figure 6.2(a)-6.2(c) shows that even small errors of predicted P_{vap} of acetone and water may result in significant uncertainty in the design. The uncertainty (in terms of average relative error) of predicted P_{vap} of acetone from DIPPR® correlation used in PRO/II® is ±0.25%. This level of accuracy of P_{vap} is found to have significant impact on the calculation of R, N and N_s (up to 20% deviation from base case, see Figure 6.2(a)-6.2(c)). If this level of uncertainty in R, N and N_s is not acceptable then user needs to collect a higher accuracy data for P_{vap} and use this data in the regression to obtain improved model for P_{vap} with desired accuracy level (that is, error less than ±0.25%). The uncertainty of predicted $\Delta_{vap}H$, ρ, and σ of acetone from DIPPR correlation used in PRO/II® is ±3.44%, ±0.38%, and ±1.77% respectively. The uncertainty of design variables R, N, N_s, Q_r and D due to given uncertainties in predicted P_{vap}, $\Delta_{vap}H$, and ρ is 19.0%, 17.0%, 18%, 5.2% and 0.50% respectively. It can be seen from Figure 6.2(d) that accurate predictions of $\Delta_{vap}H$ of acetone is important to reduce the uncertainty in calculation of Q_r. The sensitivities of D to prediction errors in ρ and σ are shown in Figures 6.2(e)-6.2(f). In this case study, the given uncertainty of ρ and σ is found to have little impact on the calculation of D. Figure 6.2(g) shows the effects of errors of critical temperature of acetone and water on the calculated reflux ratio (when SRK equation-of-state is used to model the vapor phase) and it can be seen that the given uncertainty of critical temperature of acetone (±0.4%) and water (±0.3%) has little impact on the reflux ratio. From Figures 6.2(a)-6.2(g), it can be concluded that the most sensitive properties are P_{vap} of acetone and water, and $\Delta_{vap}H$ of acetone. This information serves as input for a process design tool to carry out further improvements.

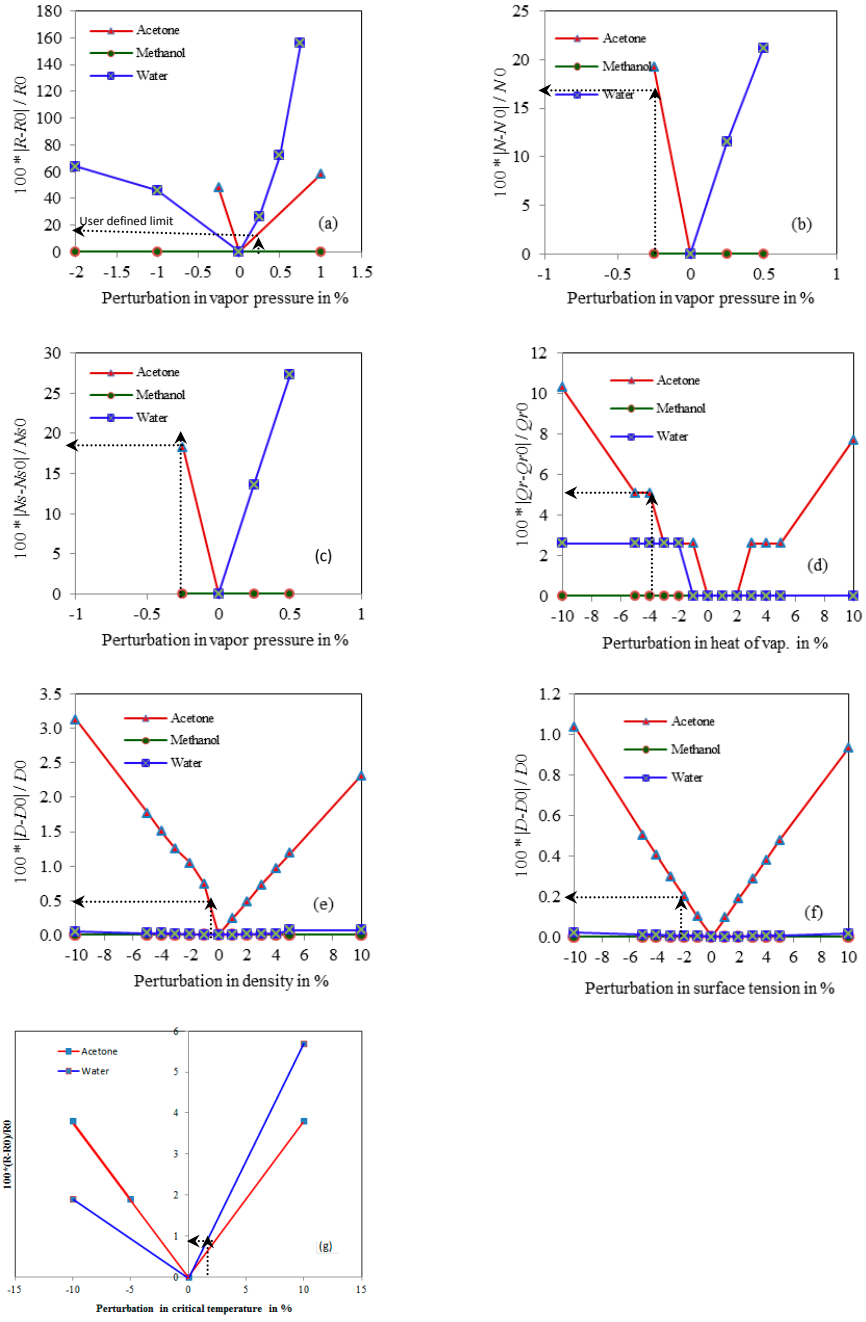


Figure 6.2. (a) Sensitivity of R due to errors of P_{vap}; (b) Sensitivity of N due to errors of P_{vap}; (c) Sensitivity of Ns due to errors of P_{vap}; (d) Sensitivity of Q_r due to errors of $\Delta_{vap}H$; (e) and (f) Sensitivity of D due to errors of ρ and σ resp.; (g) Sensitivity of R due to errors of T_c

6.2 Case Study 2: Sensitivity Analysis of Design of Short-Path Evaporator due to Uncertainties of Property Estimates

Conventional distillation process is not recommended for pure components that can be degraded under distillation temperatures, such as vitamins, insecticides, drugs and flavours/fragrances. Short-path evaporation is a separation technique used as an alternative in various processes of the chemical, pharmaceutical, fragrance and food industry. It is a safe method suitable for separation and purification of thermally unstable materials, through a small distance between the evaporator and the condenser. The short-path evaporator is characterized by low temperatures, short residence times of the distilled liquid on the thermally exposed surface, and sufficiently low pressure in the distillation gap (space between evaporator and condenser). This method is also called molecular distillation because the vapor path is unobstructed, and the condenser is separated from the evaporator by a distance less than the mean free path of the evaporating molecule (Sales-Cruz and Gani, 2006).

Step 1. Problem definition

It is required to perform sensitivity analysis of process design of short-path evaporator employed to separate glycerol from the mixture of mono-, di-, and tri-Caprylyn to obtain a product purity of at least 90.0% glycerol, subject to uncertainties of property estimates.

Step 2. Collect base case data

The base case design data is collected from Sales-Cruz and Gani (2006) and is summarized in Table 6.3. The pure component properties involved in the process design calculations are: (i) P_{vap}, (ii) T_c, (iii) P_c, (iv) ω, (v) $\Delta_{vap}H$, (vi) ρ, (vii) μ, (viii) thermal conductivity (k), and (ix) liquid heat capacity (C_p). The ρ is estimated using modified Racket equation which requires input information of T_c, P_c, and ω of pure components. The $\Delta_{vap}H$ is estimated using the input information of T_c of pure components. The necessary primary properties of mono-, di-, and tri-Caprylyn together with their uncertainty estimates are obtained by using the developed property models in this work. The experimentally measured values of necessary primary properties of glycerol and the information of uncertainties in property values are taken from the NIST® database. The uncertainty in experimental value of T_c (850 K) of glycerol is ± 4 K. The uncertainty in experimental value of P_c (75 bar) of glycerol is ± 2 bar. The ω for glycerol is predicted using the developed property model in this work and this value is 1.14 with uncertainty (likely prediction error) of ±0.03.The uncertainty in terms of average relative error of predicted P_{vap} of glycerol using the DIPPR® correlation is ±3.96%.

Table 6.3 Problem data for case study of short-path evaporator (Sales-Cruz and Gani, 2006)

Variable	Notation	Value	Units
Evaporator radius	R	0.36	M
Distance between evaporator – condenser	d	0.065	M
Evaporator length	L	3.0	M
Molecular weight [a]	MW	92.09[1]	Kg mol^{-1}
		218.29[2]	
		344.48[3]	
		470.68[4]	
Activity coefficient [a]	γ	0.9872[1]	
		0.9547[2]	
		0.9967[3]	
		0.9847[4]	
Number of intermolecular collisions	n	5	
Mean path of vapour	β	0.065	M
Reference pressure	P_{ref}	101325	Pa
Wall temperature	T_w	473	K
System pressure	P	10	Pa
Feed temperature	T_f	353	K
Feed composition [a]	$C_{i,o}$	3.555[1]	Kmol m^{-3}
		0.9708[2]	
		0.1618[3]	
		0.3236[4]	
Feed flow rates [a]	$I_{i,o}$	0.055[1]	mol s^{-1}
		0.015[2]	
		0.025[3]	
		0.005[4]	

[a] Note: 1 refers to glycerol; 2 refers to mono-caprylyn; 3 refers to di-caprylyn; and 4 refers to tri-caprylyn.

Step 3. Construct process model

A brief description of the steady-state process model developed by Sales-Cruz and Gani (2006) is given in Appendix E. This process model for short-path evaporator has been implemented by Sales-Cruz and Gani (2006) in MoT and is available in the ICAS®. In this work, the process model available in MoT is used for performing the sensitivity analysis.

Step 4. Solve, verify, and validate model with known data

The process model is solved in MoT and the obtained results are compared and validated against the results given by Sales-Cruz and Gani (2006). The obtained flow rate of glycerol of 0.052 mol/s and the product purity of glycerol of 90.5% matches well with those reported by Sales-Cruz and Gani (2006).

Step 5. Perform sensitivity analysis

The design variable considered for performing sensitivity analysis is the product flow rate containing glycerol of purity of 90.5%. The sensitivity analysis is carried out using the sensitivity analysis feature available in MoT which allows the user to select the response variables (in this case, product flow rate) and sensitivity parameters (in this case, pure component properties) namely, P_{vap}, T_c, P_c, ω, $\Delta_{vap}H$, ρ, μ, k, and C_p to be perturbed. The results of sensitivities of product flow rate due to uncertainties of property estimates in the form of plots as obtained from MoT are shown in Figures 6.3(a)-6.3(h). In these figures, the y-axis represents the deviation of the product flow rate from the base case design product flow rate. The zero point on the x-axis represents the base case design.

Step 6. Evaluation of results of sensitivity analysis

Figure 6.3(a) shows the effects of uncertainties of predicted property values of P_{vap} of glycerol and mono-, di-, and tri-caprylyn on the product flow rate from the short path evaporator. It can be seen that, the given level of accuracy of correlation for P_{vap} of glycerol may have significant impact (up to 15% deviation) on the final product flow rate.

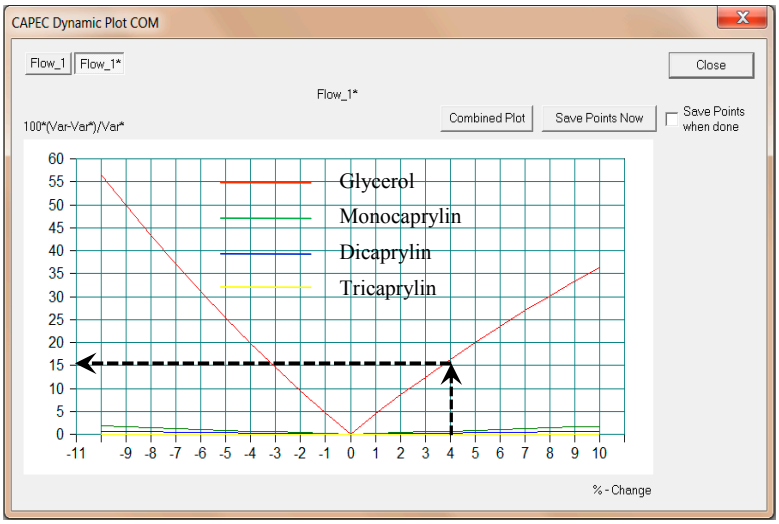


Figure 6.3(a) Sensitivity of product flow rate due to uncertainties of predicted P_{vap}

It can be seen from Figures 6.3(b)-6.3(d) that the uncertainties of predicted P_c and ω of di-caprylyn have significant impact on the product flow rate. For di-caprylyn, the experimentally measured values of T_c, P_c, and ω are not available in the literature. Hence, these properties are predicted using the developed property models in this work. The predicted T_c, P_c, and ω of di-caprylyn together with the uncertainties of predicted property values are taken from the ProPred.

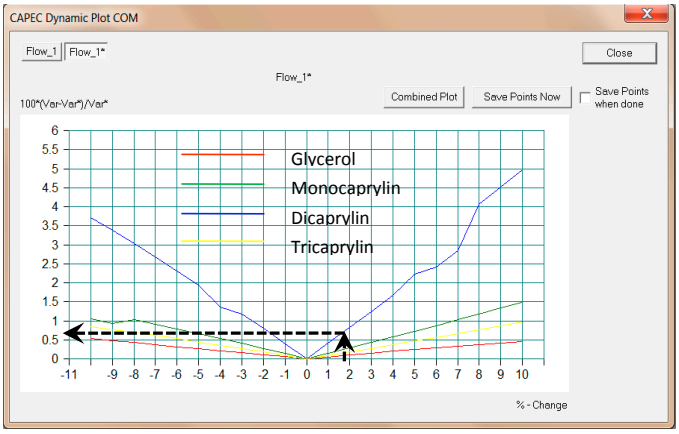


Figure 6.3(b) Sensitivity of product flow rate due to uncertainties of T_c

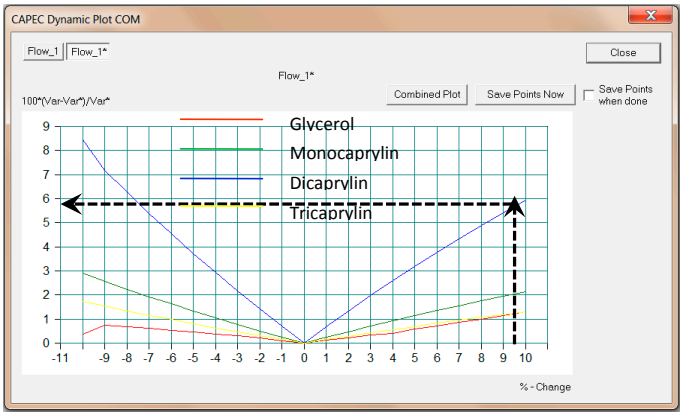


Figure 6.3(c) Sensitivity of product flow rate due to uncertainties of predicted P_c

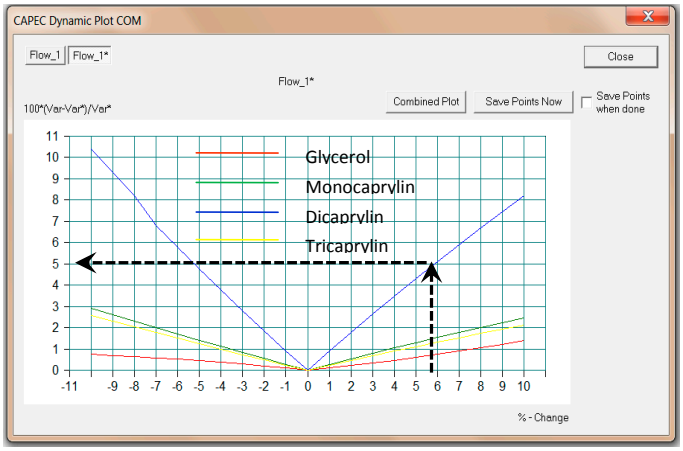


Figure 6.3(d) Sensitivity of product flow rate due to uncertainties of predicted ω

The uncertainty of predicted value of T_c (802.4 K) of dicaprylyn is ±11.82 K (±1.47%). The uncertainty of predicted value of P_c (11.40 bar) of dicaprylyn is ±1.07 bar (±9.4%). The uncertainty of predicted value of ω (1.5) of dicaprylyn is ±0.08 (±5.3%). The uncertainty of product flow rate due to given uncertainties in predicted T_c, P_c, and ω of dicaprylyn are 0.6%, 5.8%, and 5.0%, respectively. Figures 6.3(e)-6.3(h) show the effects of uncertainties of predicted C_p, k, ρ, and μ of glycerol and mono-, di-, and tricaprylyn on the product flow rate. The uncertainties of predicted C_p and k of glycerol is found to have significant impact (large slope means large uncertainty in the design variable) on the product flow rate. Both C_p and k of glycerol are important in the development of the falling film at the surface of the evaporator and the condenser. Any uncertainty in the values of C_p and k may affect the thickness of the falling film and consequently the temperature profile along the length of evaporator resulting into uncertainties in the flow rate of the glycerol in the product stream. Further, it can be seen from Figure 6.3(h) that uncertainty of predicted μ of glycerol and mono-, di-, and tricaprylyn has very little impact on the product flow rate. The uncertainty estimate of predicted ρ of dicaprylyn is not available and hence it is not indicated in Figure 6.3(g). Based on the results discussed above, the most important properties from the design of short-path evaporator point of view can be listed as: P_{vap} of glycerol, P_c of dicaprylyn ω of dicaprylyn, and T_c of glycerol. This information can be used as input to further improve the existing design through planning experimental measurements of the above listed properties or through selection of property models with desired level of accuracy.

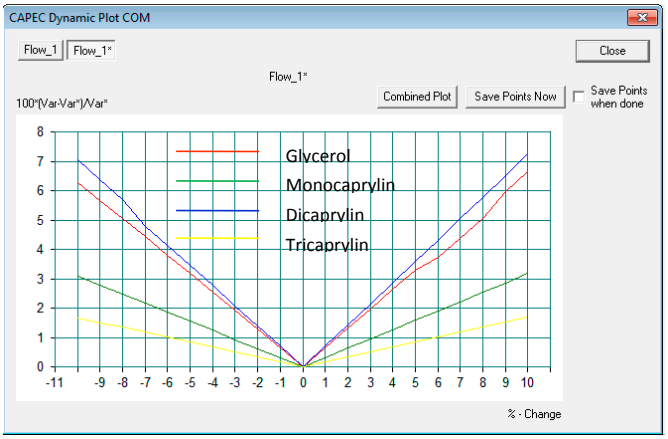


Figure 6.3(e) Sensitivity of product flow rate due to uncertainties of predicted C_p

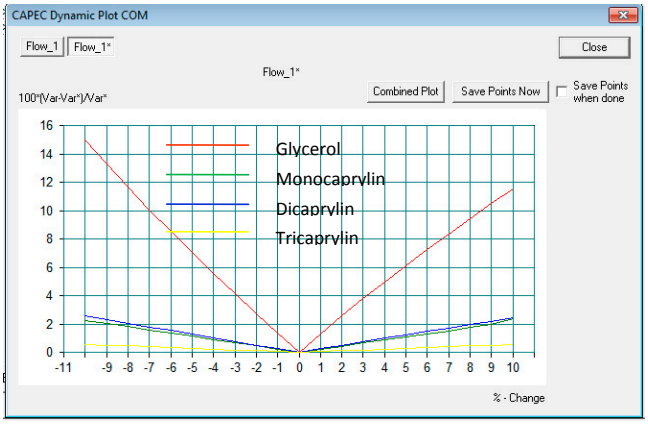


Figure 6.3(f) Sensitivity of product flow rate due to uncertainties of predicted k

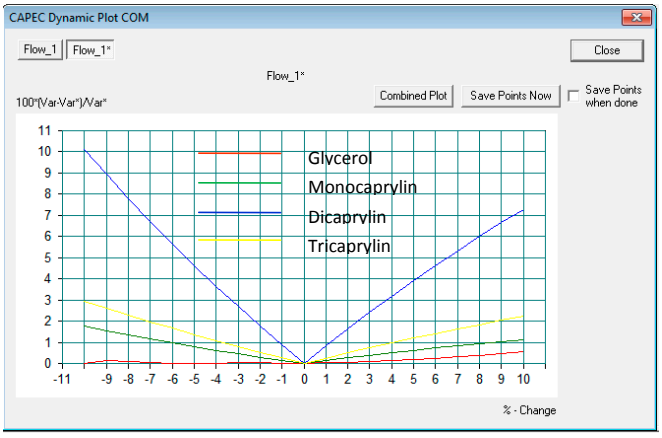


Figure 6.3(g) Sensitivity of product flow rate due to uncertainties of predicted ρ

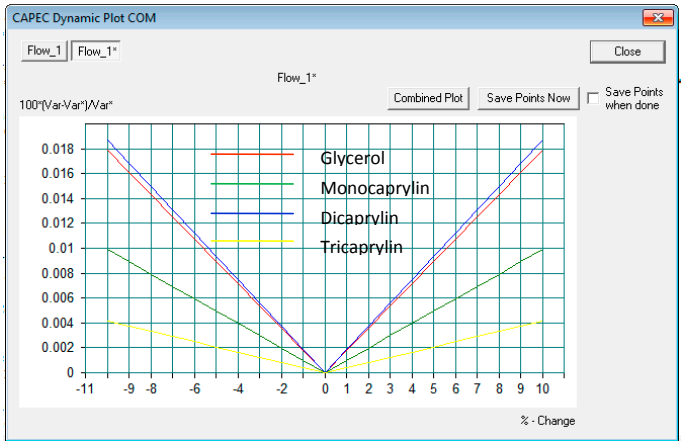


Figure 6.3(h) Sensitivity of product flow rate due to uncertainties of predicted μ

6.3. Case Study 3: Sensitivity Analysis of Process Design of De-acidification System of Deodorization Process due to Uncertainties of Property Estimates

The deacidification process is usually the final step in refining of vegetable oils. The combined steam and distilled vapors from the deodorization section are collected in a scrubber and condensed to form a distillate that can be disposed of or alternatively treated further to recover valuable byproducts (Bhosle and Subramanian, 2005). The major constituents of deodorizer distillates are FFA's, tocopherols, tocotrienols, sterols and their esters etc. which are present in various relative amounts depending on the oil sources and refining steps employed prior to the deodorization step. The commercial value of the deodorizer distillate can be greater when it is split into a FFA-enriched fraction and a fraction enriched in tocopherols, tocotrienols, sterols etc.

Step 1. Problem definition

It is required to perform sensitivity analysis of process design of a de-acidification system of crude palm oil (CPO) deodorization process due to uncertainties of property estimates. The de-acidification system is employed for the separation of FFA's and nutritional components such as tocopherols and tocotrienols (both components referred as Toco's in this work) from the vapors issuing from deodorization section. For palm oil, the de-acidification process is mostly used to increase the purity of the PFAD and to recover neutral oil (triglycerides, diglycerides, and monoglycerides) present in the distillate. The composition of FFA's in final palm oil fatty acid distillate (PFAD) from final scrubber should be at least 98%.

Step 2. Collect base case data

The composition of CPO is taken from Gee (2007) and is summarized in the Table 6.4(a) – 6.4(b). The basis and operating conditions selected for designing the de-acidification process are as below:

(1) CPO feed rate: 50,000 kg/hr
(2) Stripping steam pressure: 3 bar (abs.)
(3) Stripping steam consumption: 0.7% of CPO feed rate
(4) Vacuum maintained at the suction of vacuum system: 2 mbar
(5) Temperature of CPO entering deodorization section: 265 ^{o}C
(6) The primary and temperature dependent properties of pure components involved in the de-acidification process have been taken from the CAPEC_Lipids_Database (version 5.8) developed at CAPEC, DTU and Alfa Laval Copenhagen A/S.
(7) The UNIFAC model is employed to calculate the liquid phase activity coefficients while the vapor phase is assumed to behave as an ideal gas.

Table 6.4(a) Composition of acylglycerides in CPO (Gee, 2007)[a]

TAG's	Composition (%)	DAG's	Composition (%)	MAG's	Composition (%)
POP	42.46	P-P-OH	1.89	P-OH-OH	0.17436
PLIP	11.57	P-O-OH	3.74	O-OH-OH	0.10064
POO	24.32	P-LI-OH	0.86		
POLI	11.14				
Total	89.49%	Total	6.5%	Total	0.275%

Table 6.4(b) Composition of FFA's and other components in CPO (Gee, 2007)[a]

Others	Composition (%)	FFA's	Composition (%)
Alfa-Tocopherol	0.0218	Lauric acid	0.03
Alfa-Tocotrienol	0.0168	Myristic acid	0.052
Beta-Tocotrienol	0.0021	Palmitic acid	1.61
Gamma-Tocotrienol	0.0302	Stearic acid	0.14
Delta-Tocotrienol	0.0092	Oleic acid	1.29
Beta-Carotene	0.0369	Linoleic acid	0.31
Beta-Sitosterol	0.0219		
Campesterol	0.0088		
Stigmasterol	0.0045		
Sterol-ester	0.0040		
Cholesterol	0.0007		
Squalene	0.0350		
Water	0.0500		
Total	0.265%	Total	3.46%

[a]Notations: TAG: Triglycerides; DAG: Diglycerides: MAG: Monoglyceride; POP: 1,3-dihexadecanoyl-2-octadecenoyl-sn-glycerol; PLIP: 1,3-dihexadecanoyl-2-octadecadienoyl-sn-glycerol; POO: 1-hexadecanoyl-2,3-dioctadecenoyl-sn-glycerol; POLI: 1-hexadecanoyl-2-octadecenoyl-3-octadecadienoyl-sn-glycerol; P-P-OH:1,2-dihexadecanoyl-sn-glycerol; P-O-OH: 1-hexadecanoyl-2-octadecenoyl-sn-glycerol; P-OH-OH: 1-hexadecanoyl-sn-glycerol; O-OH-OH: 1-octadecenoyl-sn-glycerol.

The pure component properties involved in the design calculations are: (i) vapor pressure (P_{vap}), (ii) critical temperature (T_c), (iii) heat of vaporization ($\Delta_{vap}H$), (iv) liquid density (ρ), and (v) liquid viscosity (μ). The $\Delta_{vap}H$ is estimated using the input information of T_c of pure components.

Step 3. Construct process model

Figure 6.4 shows a steady state process model of de-acidification process of the deodorization system as constructed in the PRO/II® simulator.

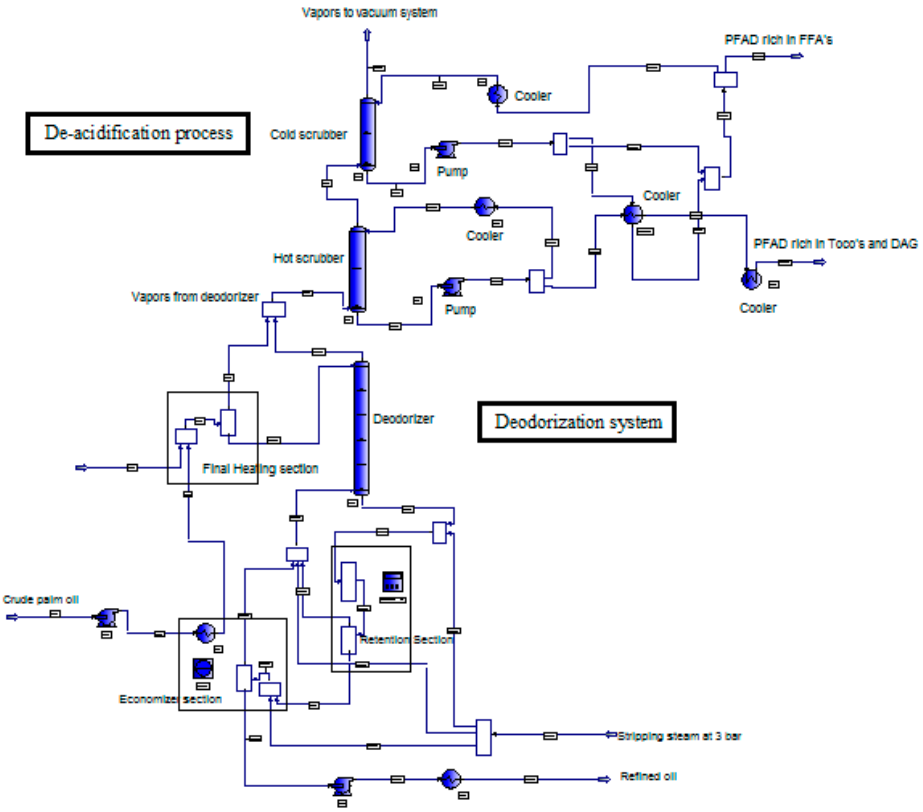


Figure 6.4 Steady state model of de-acidification process in the PRO/II®

The configuration of deodorization unit and the de-acidification process as shown in Figure 6.4 is taken from the SoftColumn^{TM} deodorization system designed and developed by Alfa Laval Copenhagen A/S. The CPO at about 100 ^{o}C is first fed to an economizer section and then it is fed to a final heater where the temperature of CPO is raised to 265 ^{o}C. The

economizer section and the final heating section are modeled as a heat exchanger combined with a flash (Ceriani et al., 2010). The configuration for these two sections in PRO/II is defined as follows. In the economizer section, the CPO flows in the tube side while stripping steam flows in the shell side of the heat exchanger. In the final heater, the CPO flashes in contact with stripping steam in the shell side of the heat exchanger. The heated CPO is then fed at the top of the deodorizer (modeled as a packed column without the reboiler and the condenser in PRO/II®). The deodorizer is divided into a continuous section in which the CPO flows downwards by gravity (most of the FFA's are removed in this section) and into a retention section where heat bleaching and deodorization takes place. The de-acidification system as shown in Figure 6.4 consists of two scrubbers with structured packing and operates under very low pressure (about 3 mbar). The vapor from the deodorizer section enters the first scrubber also called as hot scrubber. The temperature in the hot scrubber is maintained at about 180 ^{o}C and this scrubber is intended to recover valuable byproducts such as tocopherols, tocotrienols, sterols and their esters as well as to separate monoglyceride and diglycerides from the distilled vapors of deodorization system. The uncondensed vapor from hot scrubber are fed to the second scrubber also called as cold scrubber and is maintained at about 65 ^{o}C. The cold scrubber is intended to recover the FFA's from the distilled vapors.

Step 4. Solve, verify, and validate the model with known data

The process model shown in Figure 6.4 is solved in PRO/II® using the input data collected in *Step 2*. The simulated results from PRO/II ® are shown in Table 6.5 and are validated based on the industrial know-how.

Table 6.5 Stream summary for de-acidification process from PRO/II®

Step 5. Perform sensitivity analysis

Stream Description	Feed oil (CPO)	Product oil (Refined)	PFAD from Hot Scrubber	PFAD from cold scrubber
Temperature, ^{o}C	100	40	60	65
Enthalpy, kcal/hr	2.46	0.71	0.02	0.05
Molecular Weight	744.9	818.4	347.1	266.8
Total Mass Rate,	50 000	47739.5	998.7	1238.5
Composition (wt %)				
TAG's	89.50	93.53	3.31	0.00
DAG's	6.50	6.29	29.65	0.02
MAG's	0.27	0.00	11.44	1.85
FFA's	3.46	0.04	51.88	98.04
TOCO's	0.08	0.04	1.90	0.02
STEROLS	0.04	0.03	0.39	0.01
CAROTENOIDS	0.06	0.06	0.03	0.00
SQUALENE	0.03	0.01	1.39	0.06
WATER	0.05	0.00	0.00	0.00

The design variables considered for sensitivity analysis are: (i) PFAD flow from the cold scrubber; (ii) Composition of Toco's in the PFAD from hot scrubber; (iii) diameter of the hot scrubber and the cold scrubber; and (iv) the unrecovered FFA's discharged into vacuum system. The results of sensitivity analysis for selected design variables with respect to uncertainties of property estimates are shown in Figures 6.5(a)-6.5(f).

Step 6. Evaluation of results of sensitivity analysis

Figure 6.5(a) shows the effects of uncertainties of predicted P_{vap} on the PFAD flow obtained from the cold scrubber. The uncertainty in terms of average relative error of predicted P_{vap} of palmitic acid and oleic acid estimated using the DIPPR® correlations are ±4.19% and ±2.4% respectively. At these levels of uncertainties, the deviation in PFAD flow from cold scrubber can be up to 0.8%. The positive errors of predicted P_{vap} of palmitic acid and oleic acid results in lower PFAD flow rate from the cold scrubber. This implies that larger quantity of FFA's (as compared to the base case value) enters into the vacuum system which is undesirable from the recovery of FFA's point of view. This is illustrated in Figure 6.5(b) which shows effects of uncertainties of predicted P_{vap} on the quantity of FFA's discharged into the vacuum system. The negative errors of predicted P_{vap} of palmitic acid and oleic acid, on the other hand, results in higher concentration of FFA's in the PFAD from the hot scrubber which in turn causes lower composition of Toco's in the PFAD stream. This situation is illustrated in Figure 6.5(c) which shows that the given uncertainty of predicted P_{vap} of palmitic acid may result in a much lower composition of Toco's (about 1% lower). The lower composition of Toco's has direct impact on the economics of recovery of Toco's from the PFAD received from the hot scrubber.

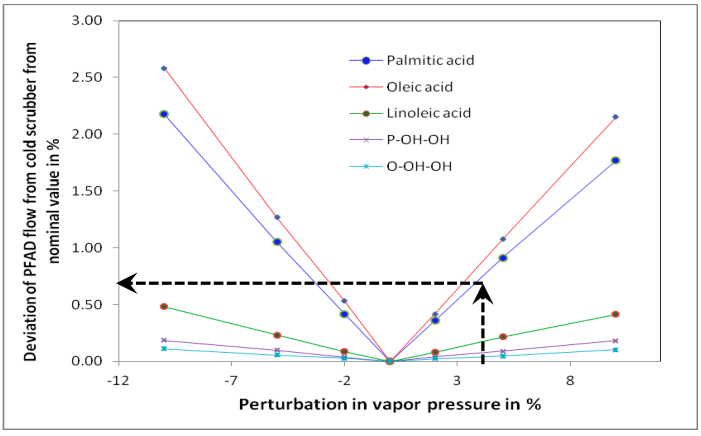


Figure 6.5(a) Sensitivities of PFAD flow from cold scrubber to uncertainties of predicted P_{vap}

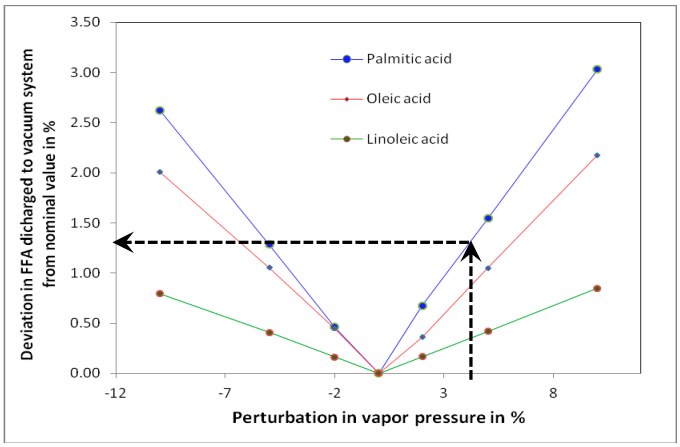


Figure 6.5(b) Sensitivities of FFA's to vacuum system due to uncertainties of predicted P_{vap}

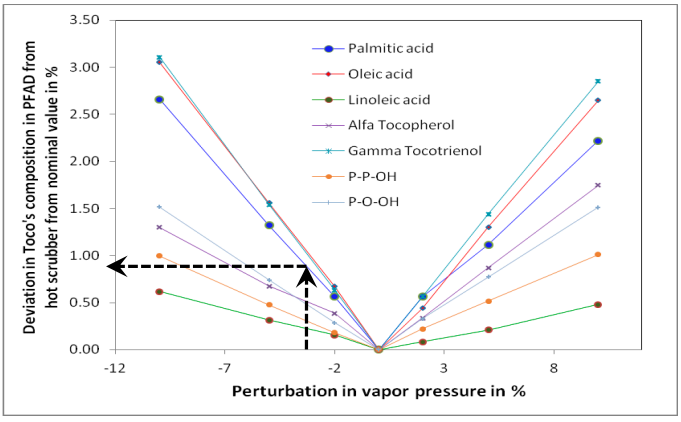


Figure 6.5(c) Sensitivities of composition of Toco's to uncertainties of predicted P_{vap}

Figure 6.5(d) and 6.5(e) shows the effects of uncertainties of predicted ρ on the diameter of hot scrubber and cold scrubber. The uncertainty in terms of average relative error of predicted ρ of palmitic acid and oleic acid estimated using the DIPPR® correlations are ±0.24% and ±0.76% respectively. At these levels of uncertainties, the sensitivity of diameter of hot scrubber and cold scrubber is insignificant. The measurement uncertainty of T_c of palmitic acid as noted from NIST® database is ±0.3%. It has been found that this level of uncertainty has little impact on the calculated PFAD flow rate from the cold scrubber (see Figure 6.5(f)). Based on the results of the sensitivity analysis, the most important properties from the design of de-acidification process of deodorization system can be listed as: P_{vap} of palmitic acid, oleic acid, linoleic acid, and gamma-tocotrienol.

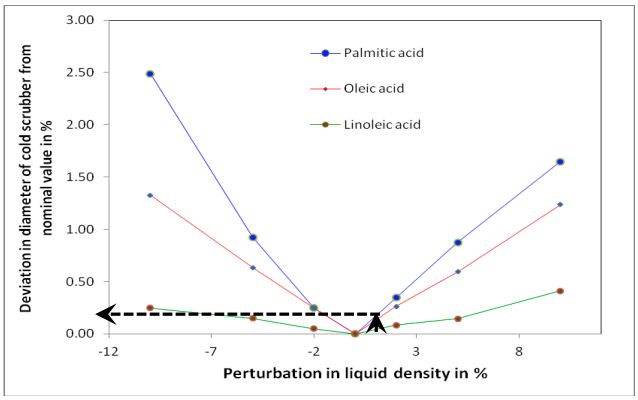


Figure 6.5(e) Sensitivity of diameter of the cold scrubber to uncertainties of predicted ρ

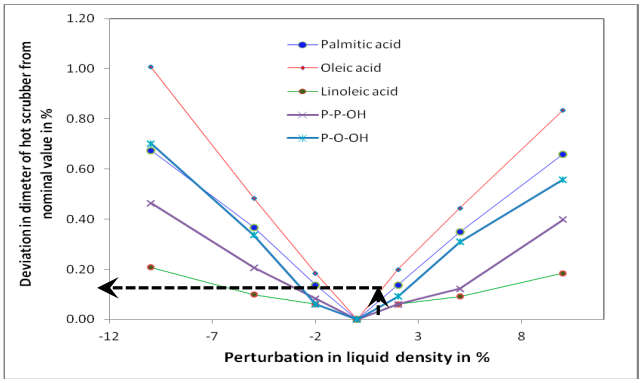


Figure 6.5(f) Sensitivity of diameter of the hot scrubber to uncertainties of predicted ρ

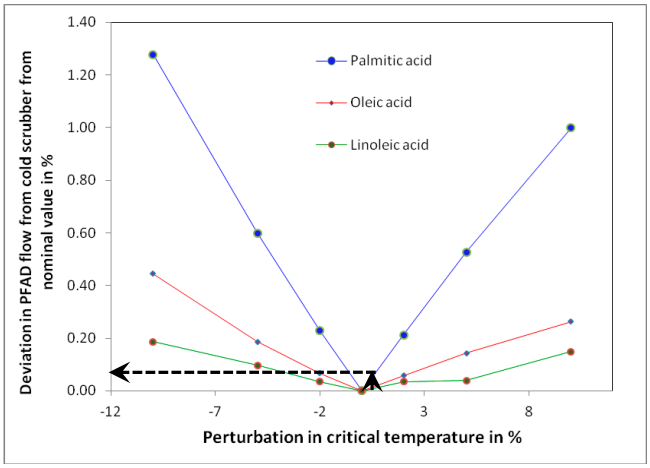


Figure 6.5(g) Sensitivities of PFAD flow from the cold scrubber to uncertainties of T_c

6.4 Conclusions

The application of the developed methodology to analyze the sensitivities of process design due to uncertainties in property estimates has been discussed through three case studies. When necessary, the developed property prediction models in this work are employed to obtain properties of pure components together with uncertainties of predicted property values. For temperature dependent properties involved in the process model, the property estimates and the related uncertainty information is obtained from the DIPPR® database. It can be observed from these three case studies that depending on the type of the process, operating conditions, and pure component properties involved, the uncertainties of the design variables due to uncertainties of pure component property estimates can be significant. In this context, sensitivity analysis of process design due to uncertainties of property estimates allows the design engineer to analyze the impact of these uncertainties on the quality (reliability) of the design and to further improve the existing design by taking necessary corrective measures. The corrective measures may include planning additional experiments to reduce the critical uncertainties of properties that are most improtant from process design point of view and the selection of appropriate safety factors in the process design.

Chapter 7. Conclusions and Future Work

7.1 Conclusions

In this work, development and analysis of GC^+ approach based property models for the estimation of pure component properties has been carried out with the aim of providing accurate and reliable estimates of pure component properties together with uncertainties of predicted property values. To this end, systematic methods for performing property modeling and uncertainty analysis have been developed and used. These methods assist property model developer in selecting the quantity of the property data-set, selecting an appropriate model function, defining structural parameters for the property model using 'molecular structural similarity criteria', and in performing tests for verification of thermodynamic consistency and predictive power of property models. These methods have been applied in the development of new property models as well as in the performance improvement of earlier versions of GC^+ based property models.

Motivated by the need of accurate and reliable predictions of thermo-physical and transport-related properties of pure components and their uncertainty estimates in chemical product and process design, the first part of the project work has been dedicated to performing property modeling and uncertainty analysis of 21 thermo-physical and transport-related properties (see chapter 4). The performance and application range of earlier version of GC^+ based property models have been improved by making use of new experimental data-points available in the extended version of the CAPEC database. Furthermore, a new method based on the simultaneous regression approach has been developed and used for determining the values of GC model parameters in a single step. The reported range of experimental measurement uncertainties for the properties is collected from the DIPPR® database and is compared with the accuracy and reliability of developed GC models. In addition, tests for verifying thermodynamic consistency and predictive power of the developed GC models have been performed. The behavior of critical constants and normal boiling point of pure components predicted using the developed GC models has been found consistent with the theoretical foundations discussed by Tsonopoulos (1987). The most important feature of the developed GC^+ based property models is that the user can obtain accurate and reliable predictions of pure component properties together with the estimates of uncertainties (in terms of 95% confidence intervals) of predicted property values. This information allow the chemist/ process engineer to take into account the effects of uncertainties of property estimates in the process and product design calculations thereby contributing to better-informed and reliable design solutions.

Another important area is the synthesis, design, and analysis of sustainable chemical products and processes where design engineers/health, safety and environment (HSE) specialists frequently find the need for estimating necessary environmental-related properties of chemicals. The range of these environmental-related properties is very large which includes, LC_{50}, aqueous solubility, bio-concentration factor, LD_{50}, photochemical oxidation potential, global warming potential, emission to urban air (carcinogenic and non-carcinogenic) etc. to name a few. Since there are limited numbers of reliable property models available for the estimation of these important properties, the focus of second part of this work has been concentrated on the development and analysis of GC^{+} based property models to provide reliable estimation of 22 environmental-related properties of pure components (see chapter 4) together with uncertainty estimates. The developed methods employed for modeling and analysis of thermo-physical properties are also used for modeling and analysis of environmental-related properties. To highlight the importance of considering all of the data-points in the regression, cross validation analysis for LC_{50}, LD_{50}, and emission to urban air (carcinogenic and non-carcinogenic) has been performed. For the case when it is preferred to retain some of the data-points for the model validation purpose, a method for selecting minimum data-set for the parameter regression is developed.

For some properties of pure components such as enthalpy of formation, it is desirable to have property prediction models that can yield a higher level of prediction accuracy (for example to reach a chemical accuracy of ± 2 kJ/mol). To achieve the desired level of prediction accuracy, a method based on the "molecular structure similarity criteria" is developed. This method has been successfully applied to improve the prediction performance of GC based models for the estimation of pure components properties namely, enthalpy of formation, enthalpy of fusion, and critical temperature.

The developed GC^{+} based property models for the estimation of thermo-physical, environment-related and transport-related properties have been implemented in the ProPred software with the aim of providing the user a fast way of estimation of needed pure component properties. Finally, the effects of uncertainties of predicted pure component values on the quality of design have been analysed through application of developed method for performing sensitivity analysis of process design due to uncertainties of property estimates – three case studies were studied to this end. Such analysis helps the process design engineers to identify the most critical properties from design point of view and to take corrective measures (for example, plan necessary experimental measurements, safety factors etc.) in order to reduce the uncertainties in the final design and deliver robust solutions.

7.2 Achievements and Contributions

The achievements and contributions from this PhD project can be summarized as follows:

1. Development of systematic methods for performing property modeling and uncertainty analysis of GC+ based property models for the estimation of pure component properties. The application of these methods allows one to obtain predicted property values together with uncertainties of predicted property values -much needed information for assessing the quality of the product and process design.

2. Successful application of developed methods for property modeling and uncertainty analysis in modeling and analysis of 21 thermo-physical and transport-related properties and 22 environment-related properties of pure components.

3. Development of 'molecular structural similarity criteria' based approach for developing GC based property models to guide the model developers systematically to reach a higher target of prediction accuracy.

4. The GC based property model for enthalpy of formation developed using the 'molecular structural similarity criteria' approach yields average absolute deviation of just 1.75 kJ/mol which is well within the desired chemical accuracy of ±2%. The new third-order group definitions developed for the enthalpy of formation property model were evaluated and verified successfully for modeling of heat of fusion and critical temperature properties. The method to identify the new third-order groups in the molecular structure of pure components needs to be developed.

5. Implementation of developed property prediction models into the ProPred software for easy and fast estimation of pure component properties.

6. Development of a systematic method for performing sensitivity analysis of process design due to uncertainties of property estimates. This method allows the engineer to obtain quantitative and useful information related to potential impact that the uncertainties of property estimates would have on the quality of final design.

Last but not least, the resulting list of publications as well as the conference presentations related to this PhD project is presented in Appendix F.

7.3 Future Work and Recommendations

As noted above**,** the methods related to property modeling and uncertainty analysis have been successfully developed and applied to modeling of wide range of pure component properties. There are, however, further opportunities for increasing the scope of these methods. In this context, following are the most important suggestions/recommendations for the future work.

1. Application of the 'molecular structural similarity criteria' based approach is recommended for improving the existing performance of GC models of other important pure component properties such as normal boiling point, normal melting point etc.

2. Extension and testing of the developed method for finding minimum data-set for the parameter regression is recommended for modeling other pure component properties. For extension of the method, concepts from DoE (design of experiments) statistical framework with proper design criteria such as D-optimality which is relevant for parameter estimation accuracy can be explored.

3. The application of developed methodology for performing sensitivity analysis of process design due to uncertainties of property estimates is recommended for other important areas such as chemical product design (for example, selection and design of solvents), sustainable process design etc. In addition, the methodology can be extended to provide an overall risk of failure for a given process design under a given set of uncertainties that goes into a design efforts. To this end, uncertainty analysis can be added to the methodology.

4. The extension of developed methods for property modeling and uncertainty analysis to analyze temperature dependent properties of pure components such as, liquid vapor pressure, heat of vaporization, liquid density etc.

5. The effect of uncertainties of experimentally measured property values on the quality of parameter estimation needs to be investigated by considering these uncertainties as weighing factors for data-points (better quality means higher weight and more reliability) in weighted least square regression analysis.

6. For property modeling and uncertainty analysis of certain properties which do not follow additivity principle (such as, standard Gibbs free energy of formation), careful analysis of applicability of GC approach needs to be performed.

Nomenclature

AAE	average absolute error
AE_{max}	maximum absolute error found from the regression
A_i	occurrence of atom of type-*i*
a_i	contribution of atom of type-*i*
ARE	average relative error [%]
b	adjustable parameter of CI model
B (1/2, v/2)	beta function
c	adjustable parameter of CI model
CI	atom connectivity index
C_i	contribution of first-order group of type-*i*
$COV(\boldsymbol{P}^*)$	covariance matrix
C_p	specific heat [kJ/molK]
d	universal parameter of CI model
d_c	critical density [kg/m^3]
D_j	contribution of second-order group of type-*j*
E_k	contribution of third-order group of type-*k*
$f(X)$	function of property X
F_p	flash point [K]
GC	group-contribution
GC$^+$	group-contribution$^+$
$J(\boldsymbol{P}^*)$	local sensitivity of the model to variations in estimated model parameters
k	liquid thermal conductivity
$Logk_{ow}$	octanol/water partition coefficient
MG	Marrero and Gani
M_j	occurrence of second-order group of type-*j*
$MSECV$	mean squared error of cross-validation
N	number of experimental data-points used in the regression
N_i	occurrence of first-order group of type-*i*
N_s	entrainer stage
O_k	occurrence of third-order group of type-*k*
$\boldsymbol{P}$	model parameters
$\boldsymbol{P}^*$	estimated values of model parameters
P_c	critical pressure [bar]
$P_{c\infty}$	limiting value for critical pressure at very high carbon number [bar]
$P_r(t, v)$	Students t-distribution function
P_{rc}	percentage of the experimental data-points [%]
P_{vap}	vapor pressure [bar]

Q_c	condenser duty
Q_r	reboiler duty
R	reflux ratio
R^2	coefficient of determination
RE_{max}	maximum relative error found from the regression [%]
$S(\boldsymbol{P})$	cost function
SD	standard deviation
SSE	minimum sum of squared errors
$t(v, \alpha_t/2)$	t-distribution value corresponding to the $\alpha_t/2$ percentile
T_{ait}	Auto ignition temperature [K]
T_b	normal boiling point [K]
T_c	critical temperature [K]
T_m	normal melting point [K]
V_c	critical volume [cc/mol]
V_m	liquid molar volume at 298 K [cc/kmol]
X^{exp}	experimental property value
X^{pred}	predicted property value
Z_c	critical compressibility factor
$\Delta_f G^o$	standard Gibbs energy of formation [kJ/mol]
$\Delta_f H^o_{gas}$	standard enthalpy of formation [kJ/mol]
$\Delta_{fus} H$	normal enthalpy of fusion [kJ/mol]
$\Delta_{vap} S$	entropy of vaporization at the normal boiling point [J/mol-K]
$\Delta_{vap} H$	enthalpy of vaporization at the normal boiling point [kJ/mol]
$\Delta_{vap} H°$	enthalpy of vaporization at 298 K [kJ/mol]

Greek Symbols

ω	acentric factor
${}^{v}\chi^{0}$	zeroth-order (atom) connectivity index
${}^{v}\chi^{1}$	first-order (bond) connectivity index
ν	degrees of freedom
δ	Hildebrand solubility parameter [$MP^{1/2}$]
δ_D	Hansen solubility parameter-dispersion [$MP^{1/2}$]
δ_H	Hansen solubility parameter-hydrogen bond [$MP^{1/2}$]
δ_P	Hansen solubility parameter-polar [$MP^{1/2}$]
ρ	liquid density
μ	liquid viscosity [cP]
σ	liquid surface tension [N/m]
χ^2_{red}	reduced Chi-square

References

Abramowitz, M.; Stegun, I. A. Handbook of mathematical functions with formulas, graphs, and mathematical tables; New York: Dover, 1972.

Ambrose, D. Correlation and Estimation of Vapor-Liquid Critical Properties: I. Critical Temperatures of Organic Compounds. 1978, NPL Rep. Chem. 92, Nat. Physical Lab., Teddington, UK.

Benson, S. W.; Cruickshank, F. R.; Golden, D. M.; Haugen, G. R.; O'Neal, H. E.; Rodgers, A. S.; Shaw, R.; Walsh, R. Additivity rules for the estimation of thermochemical properties. *Chem. Rev*. **1969**, *69*, 279.

Bhosle, B. M.; Subramanian, R. New approaches in deacidification of edible oils: a review. J. Food Eng. 2005, 69, 481.

Boese, A. D.; Oren, M.; Atasoylu, O; Martin, J. M. L; Kállay, M. W3 theory: Robust computational thermochemistry in the kJ/mol accuracy range. *Journal of Chemical Physics*. **2004**, *120*, 4129-4141.

Boethling, R. S.; Howard, P. H.; Meylan, W. M. Finding and estimating chemical property data for environmental assessment. *Environ. Toxicol. Chem.* **2004**, *23*, 2290-2308.

Brignole, E. A.; Gani, R.; Romagnoli, J. A. A simple algorithm for sensitivity and operability analysis of separation processes. *Ind. Eng. Chem. Process. Des. Devel*. **1985**, *24*, 42–48.

Bunz, A .P.; Braun, B.; Janowsky, R. Quantitative structure–property relationships and neural networks: correlation and prediction of physical properties of pure components and mixtures from molecular structure, *Fluid Phase Equilib*. 1999. *158–160*, 367–374.

Casalegno, M.; Benfenati, E.; Sello, G. An automated group contribution method in predicting aquatic toxicity: the diatomic fragment approach. *Chem. Res. Toxicol*. **2005**, *18*, 740-746.

Chan, B.; Deng, J.; Radom, L. G4 (MP2)-6x: A cost-effective improvement to G4 (MP2). *J. Chem. Theory Comput.* **2011**. *7*, 112-120.

CHETAH- ASTM Computer Program for Chemical Thermodynamic and Energy Release Evaluation—CHETAH Version 9.0, http://www.normas.com/ASTM/COMPs/CHETAH-9-0.html (accessed March 18, 2013).

Ceriani,R.; Gani, R.; Meirelles, A.J.A. Simulation of thin-film deodorizers in palm oil refining. *J Food Process Eng*. **2010**, *33*, 208-225.

Chickos, J. S.; Acree, W. E., Jr. Total Phase Change Entropies and Enthalpies: An Update on Fusion Enthalpies and Their Estimation. *Thermochim. Acta* **2009**, *495*, 5–13.

Cohen, N.; Benson, S. W. Estimation of heats of formation of organic compounds by additivity methods. *Chem. Rev*. **1993**, *93*, 2419.

Constantinou, L.; Gani, R. New Group-Contribution Method for the Estimation of Properties Of Pure Compounds. *AIChE J*. **1994**, *10*, 1697-1710.

Constantinou, L.; Gani, R.; O'Connell, J. P. Estimation of the acentric factor and the liquid molar volume at 298K using a new group contribution method. *Fluid Phase Equilib*. **1995**, *103*, 11–22.

Conte, E.; Martinho, A.; Matos, H. A.; Gani, R. Combined Group-Contribution and Atom Connectivity Index-Based Methods for Estimation of Surface Tension and Viscosity. *Ind. Eng. Chem. Res*. **2008**, *47*, 7940–7954.

Cunico, L. P., Hukkerikar, A. S., Ceriani, R., Sarup, B., and Gani, R. Molecular structure-based methods of property prediction in application to lipids: A review and refinement. *Fluid Phase Equilibria*. 2013 (In Press).

Curtiss, L. A.; Raghavchari, K.; Trucks, G. W; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) theory for molecules containing first and second-row atoms. *Journal of Chemical Physics*. **1998**, *109*, 7764-7776.

Curtiss, L. A.; Raghavchari, K.; Trucks, G. W; Pople, J. A. Gaussian-2 theory for molecular energies of first-and second-row compounds. J*ournal of Chemical Physics*. **1991**, *9*4, 7221-7230.

Dalmazzone, D.; Salmon, A.; Guella, S. A second order group contribution method for the prediction of critical temperatures and enthalpies of vaporization of organic compounds. *Fluid Phase Equilib*. **2006**, *242*, 29–42.

Diaz-Tovar, C., Mustaffa, A., Hukkerikar, A.S, Quaglia, A., Sin, G., Kontogeorgis, G, Sarup, B. Gani, R. Lipid Processing Technology: Building a Multilevel Modeling Network. *Computer Aided Chemical Engineering*. 2011, 29, 256-260.

Dohrn, R.; Pfohl, O. Thermophysical properties-Industrial directions. *Fluid Phase Equilib*. **2002**, *194-197*, 15.

Domalski, E. S.; Hearing, E. D. Estimation of the Thermodynamic Properties of C-H-N-O-S-Halogen Compounds at 298.15 K. *J. Phys. Chem. Ref. Data* **1993**, *22*, 805.

Frenkel, M.; Chirico, R. D.; Diky, V.; Yan, X. J.; Dong, Q.; Muzny, C. ThermoData Engine (TDE): Software implementation of the dynamic data evaluation concept. *J. Chem. Inf. Model*. **2005**, *45* (4), 816–838.

Gani, R.; Constantinou, L. Molecular structure based estimation of properties for process design. *Fluid Phase Equilib*. **1996,** *116*, 75-86.

Gani, R.; Harper, P. M.; Hostrup, M. Automatic creation of missing groups through connectivity index for pure-component property prediction. *Ind. Eng. Chem. Res.* **2005**, *44*, 7262-7269.

Gee, P. T. Analytical characteristics of crude and refined palm oil and fractions. *European Journal of Lipid Science and Technology*. **2007**, *109*, 373-379.
Gelman A.; Carlin J.B.; Stern H.S.; Rubin D.B. Bayesian Data Analysis. 1995. Chapman and Hall, London.

Gil, I. D.; Botı´a, D. C.; Ortiz, P.; Sa´nchez, O. F. Extractive Distillation of Acetone/Methanol Mixture Using Water as Entrainer. *Ind. Eng. Chem. Res*. **2009**, *48*, 4858.

Hajipour, S.; Satyro, M. A. Uncertainty Analysis Applied to Thermodynamic Models and Process Design − 1. Pure Components. *Fluid Phase Equilib*. **2011**, *307*, 78−94.

Hall, L. H.; Kier, L. B. Molecular connectivity in chemistry and drug research; Academic Press: New York, 1976.

Heijungs, R.; Guinée, J.; Huppes, G.; Lankreijer, R. M.; Udo de Haes, H. A.; Wegener, A.; Sleeswijk, A. M. M.; Ansems, P. G.; Eggels, R.; van Duin; de Goede, H. P. Environmental Life Cycle Assessment of Products. Guide and Backgrounds; CML (Center of Environmental Science), Leiden, 1992.

Hukkerikar, A. S., Sarup, B., Ten Kate, A., Abildskov, J., Sin, G., and Gani, R. Group contribution^{+} (GC^{+}) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilibria*. 2012, 321, 25-43.

Hukkerikar, A. S., Kalakul, S., Sarup, B., Young, D. M., Sin, G., and Gani, R. Estimation of environment-related properties of chemicals for design of sustainable processes: Development of group-contribution^{+} (GC^{+}) models and uncertainty analysis. *Journal of Chemical Information and Modeling.* 2012, 52 (11), 2823-2839.

Hukkerikar, A. S., Jones, M., Sarup, B., Abildskov, J., Sin, G., and Gani, R. Sensitivity of process design due to uncertainties in property estimates. *Computer Aided Chemical Engineering*. 2012, 31, 200-204.

Hukkerikar, A. S., Meier, R. J., Sin, G., and Gani, R. A method to estimate the enthalpy of formation of organic compounds with *chemical accuracy*, *Fluid Phase Equilibria*. 2013, 348, 23-32.

Hukkerikar, A. S., Sarup B., Sin, G., and Gani, R. Molecular structure based property modeling: Development/improvement of property models through a systematic property-data-model analysis. 2013, PPEPPD-2013 conference.

Integrated Computer Aided System (ICAS), version 16.0; Department of Chemical Engineering, Technical University of Denmark, Lyngby, Denmark, 2012.

J. C. Slater, A simplification of the Hartree-Fock method. *Physical Review*. **1951**, *81*, 385-390.

J. S. Chickos, W. E. Acree Jr., J. F. Liebman, in: Frurip D, Irikura KK (Eds.) Computational thermochemistry, prediction and estimation of molecular thermodynamics. ACS Symposium Series, ACS, 677, Washington DC, 1998, pp.63–93.

Jain, A.; Yang, G.; Yalkowsky, S. H. Estimation of Melting Points of Organic Compounds. *Ind. Eng. Chem. Res*. **2004**, *43*, 7618-7621.

Joback, K. G.; Reid, R. C. Estimation of Pure-Component Properties from Group Contributions. *Chem. Eng. Cornmun*. **1987**, *57*, 233.

Kang, J. W.; Abildskov, J.; Gani, R.; Cobas, J. Estimation of mixture properties from first and second-order group contributions with the UNIFAC model. *Ind. Eng. Chem. Res*. **2002**, *41* (13), 3260-3273.

Klincewicz, K.M.; Reid, R.C. Estimation of critical properties with group contribution methods. *AIChE* J. **1984**, *30* 137–142.

Klopman, G.; Zhu, H. Estimation of the aqueous solubility of organic molecules by the group contribution approach. *J. Chem. Inf. Model.* **2001**, *41*, 439-445.

Kolska´, Z.; Růžička, V.; Gani, R. Estimation of enthalpy of vaporization and of entropy of vaporization for pure organic compounds at 298.15 K and at normal boiling temperature by a group contribution method. *Ind. Eng. Chem. Res*. **2005**, *44*, 8436-8454.

Kontogeorgis, G. M.; Gani, R. Computer Aided Property Estimation for Process and Product Design. Computer-Aided Chemical Engineering; Elsevier; Amsterdam, 2004; Vol. 19.

Kühne, R.; Ebert, R.-U.; Kleint, F.; Scmidt, G.; Schüürmann, G. Group contribution methods to estimate water solubility of organic chemicals. *Chemosphere*. **1995**, *30*, 2061-2077.

Larsen, A. H. Data Quality for Process Design. *Fluid Phase Equilib*. **1986**, *29*, 47.

Lydersen, A. L. Estimation of Critical Properties of Organic Compounds, Eng. Exp. Stn. Rep. 3; University of Wisconsin College Engineering: Madison, WI, 1955.

Macchietto, S.; Maduabeuke, G.; Szczepanski, R. Exact Determination of Process Sensitivity to Physical Properties. *Fluid Phase Equilib*. **1986**, *29*, 59.

Maranas, C. D. Optimization accounting for property uncertainty in polymer design. *Comput. Chem. Eng*. **1997**, *21*, 1019-1024.

Marrero, J.; Gani, R. A group contribution based estimation of octanol-water partition coefficient and aqueous solubility. *Ind. Eng. Chem. Res.* **2002**, *41*, 6623-6633.

Marrero, J.; Gani, R. Group-contribution based estimation of pure component properties. *Fluid Phase Equilib.* **2001**, *183-184*, 183-208.

Marrero, J.; Pardillo, E. Estimation of pure compound properties using group-interaction contributions. *AIChE J*. **1999**, *45*, 615-622.

Martin T.; Harten, P.; Venkatapathy, R.; Das, S.; Young, D. A Hierarchical Clustering Methodology for the Estimation of Toxicity. *Toxicol. Mech. Methods*. **2008**, *18*, 251-266.

Martin, J. M. L.; Oliveira, G. Towards standard methods for benchmark quality ab initio thermochemistry—W1 and W2 theory. *Journal of Chemical Physics*. **1993**, 1843-1856.

Martin, T. M.; Young, D.M. Prediction of the acute toxicity (96-h LC50) of organic compounds to the fathead minnow (*pimephales promelas*) using a group contribution method. *Chem. Res. Toxicol*. **2000**, *14*, 1378-1385.

McFarlane, R. A.; Gray, M. R.; Shaw, J. M. Evaluation of co-volume mixing rules for bitumen liquid density and bubble pressure estimation. *Fluid Phase Equilib*. **2010**, *293*, 87.

Mevik, B. H.; Cederkvist, H. R. Mean squared error of prediction (MSEP) estimates for principal component regression (PCR) and partial least squares regression (PLSR). *J. Chemom*. **2004**, *18*, 422-429.

Modarresi, H.; Conte, E.; Abildskov, J.; Gani, R.; Crafts, P. Model based calculation of solid solubility for solvent selection - A review. *Ind. Eng. Chem. Res*. **2008**, *47*, 5234–5242.

Müller, K.; Mokrushina, L.; Arlt, W. Second-Order Group Contribution Method for the Determination of the Dipole Moment. *J. Chem. Eng. Data*. **2012**, *57*, 1231−1236.

Nagano, Y. J. Standard enthalpies of formation of phenanthrene and naphthacene. *Chem. Thermodyn*. **2002**, *34*, 377-383.

Nannoolal, Y.; Rarey, J.; Ramjugernath, D. Estimation of Pure Component Properties. Part 2. Estimation of Critical Property Data by Group Contribution. *Fluid Phase Equilib*. **2007**, *252*, 1–27.

Nannoolal, Y.; Rarey, J.; Ramjugernath, D. Estimation of pure component properties. Part 4: Estimation of the saturated liquid viscosity of non-electrolyte organic compounds via group contributions and group interactions *Fluid Phase Equilib*. **2009**, *281*, 97–119.

Nannoolal, Y.; Rarey, J.; Ramjugernath, D.; Cordes W. Estimation of pure component properties: Part 1. Estimation of the normal boiling point of non-electrolyte organic compounds via group contributions and group interactions. *Fluid Phase Equilib*. **2004**, *226*, 45.

Nielsen, T. L.; Abildskov, J.; Harper, P. M.; Papaeconomou, I.; Gani, R. The CAPEC Database. *J. Chem. Eng. Data*. **2001**, *46*, 1041–1044.

NIST Chemistry WebBook, http://webbook.nist.gov, (retrieved March 18, 2013).

OSHA PEL. http://www.osha.gov/SLTC/pel/ (accessed September 12, 2012).

Parks, G. S.; West, T. J.; Naylor, B. F.; Fujii, P. S.; McClaine, L. A. Thermal Data on Organic Compounds. XXIII. Modern Combustion Data for Fourteen Hydrocarbons and Five Polyhydroxy Alcohols. *J. Am. Chem. Soc*. **1946**, *68*, 2524–2527.

Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical accuracy in ab initio thermochemistry and spectroscopy: current strategies and future challenges. *Theor. Chem. Acc.* **2012***, 131*, 1079-1098.

Poling, B. E.; Prausnitz, J. M.; O'Connell, J. P. Properties of Gases and Liquids, 5th ed.; McGraw-Hill: New York, 2000.

Project 801, Evaluated Process Design Data, Public Release Documentation; Design Institute for Physical Properties (DIPPR), American Institute of Chemical Engineers (AIChE): Provo, UT, 2006.

Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. PDDG/PM3 and PDDG/MNDO: improved semiempirical methods. *Journal of Computational Chemistry*. **2002**, *23*, 1601-1622.

Rosenbaum R. K.; Huijbregts M. A. J.; Henderson A. D.; Margni M.; McKone T. E.; van de Meent D.; Hauschild M. Z.; Shaked S.; Li D. S.; Gold L. S.; Jolliet O. USEtox human exposure and toxicity factors for comparative assessment of toxic emissions in life cycle analysis: sensitivity to key chemical properties. *Int. J. Life Cycle Assess.* **2011**, *16*, 710–727.

Rostislavovich Shafarevich I.; Alekseĭ O. Remizov, Matrices and Determinants in Linear algebra and geometry. 2013, pp. 55-63, Heidelberg: Springer.

Roux, M. V.; Temprado, M.; Chickos, J. S.; Nagano, Y. Critically evaluated thermochemical properties of polycyclic aromatic hydrocarbons. *J. Phys. Chem. Ref. Data* **2008**, *37*, 1855.

John Robert Taylor: An introduction to error analysis, page 268. University Science Books, 1997.

Sales-Cruz, M.; Gani, R. Computer-aided modelling of short-path evaporation for chemical product purification, analysis and design. *Chem. Eng. Res. Des*. **2006**, *84*, 583.

Satayanarayana, K. C.; Gani, R.; Abildskov, J. Polymer Property Modeling Using Grid Technology for Design of Structured Products. *Fluid Phase Equilib*. **2007**, *261*, 58–63.

Satyanarayana, K. C.; Abildskov J.; Gani, R. Computer-aided polymer design using group contribution plus property models. *Comput. Chem. Eng*. **2009**, *33*, 1004–1013.

Seader, J.; Henley, M. Separation Process Principles; Wiley: New York, 1998.

Seber, G.; Wild, C. Nonlinear regression; Wiley: New York, 1989.

Simamora, P.; Yalkowsky, S. H. Group Contribution Methods for Predicting the Melting Points and Boiling Points of Aromatic Compounds. *Ind. Eng. Chem*. *Res*. **1994**, *33*, 1405-1409.

Simsci-Esscor, PRO/II, User Guide, Simsci-Esscor: Houston, TX, Available from www.simsci-esscor.com.

Sin, G.; Meyer, A. S.; Gernaey, K.V. Assessing reliability of cellulose hydrolysis models to support biofuel process design – identifiability and uncertainty analysis. *Comput. Chem. Eng.* **2010**, *34*, 1385–1392.

Tabernero, A.; Martin del Valle, E. M.; Galan, M. A. Estimation of sublimation enthalpies of solids constituted by aromatic and/or polycyclic aliphatic rings by using a group contribution method. *AIChE* J. **2012**, *58*, 2875−2884.

Tsonopoulos, C. Critical constants of normal alkanes from methane to polyethylene. *AlChE J*. **1987**, *33*, 2080-2083.

Tsonopoulos, C.; Tan, Z. The critical constants of normal alkanes from methane to polyethylene. II. Application of the Flory theory. *Fluid Phase Equilib*. **1993**, *83*, 127–138.

U.S. Environmental Protection Agency (2012b). Toxicity Estimation Software Tool (T.E.S.T.). http://www.epa.gov/nrmrl/std/qsar/qsar.html (accessed September 12, 2012).

USEtox™ model. http://www.usetox.org/ (accessed September 12, 2012).

Van Speybroeck, V.; Gani, R.; Meier, R. J. The calculation of thermodynamic properties of molecules. *Chem. Soc. Rev*. **2010**, *39*, 1764–1779.

Whiting, W. B.; Tong, T.; Reed, M. E. Effect of Uncertainties in Thermodynamic Data and Model Parameters on Calculated Process Performance. *Ind. Eng. Chem. Res*. **1993**, *32*, 1367.

Young, D. M.; Cabezas, H. Designing Sustainable Processes with Simulation: The Waste Reduction (WAR) Algorithm.*Comput. Chem. Eng*.**1999**, *23*, 1477–1491.

Appendix A

Model performance statistics, group-contribution tables and atom-contribution tables for GC$^+$ models for the estimation of thermo-physical and transport-related properties of pure components.

Table A1. Performance of MG method based property models for thermo-physical and transport-related properties analysed using simultaneous regression method

Table A2. MG method based property models analysed using step-wise regression method: First-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, ω, and V_{mo}

Table A3. MG method based property models analysed using step-wise regression method: Second-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, ω ,and V_{mo}

Table A4. MG method based property models analysed using step-wise regression method: Third-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, ω ,and V_{mo}

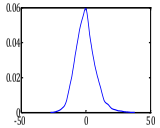
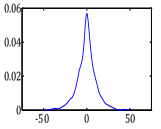
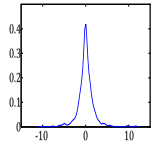
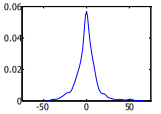
Table A5. MG method based property models analysed using simultaneous regression method: First-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, AiT, ω, and V_{mo}

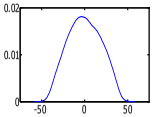
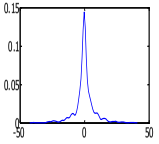
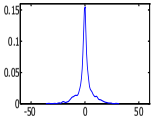
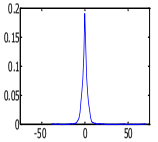
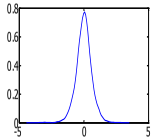
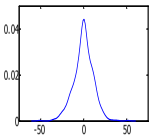
Table A6. MG method based property models analysed using simultaneous regression method: Second-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, AiT, ω, and V_{mo}

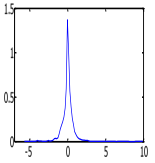
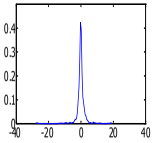
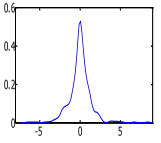
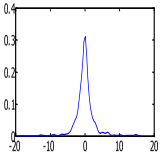
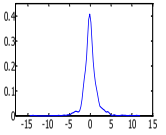
Table A7. MG method based property models analysed using simultaneous regression method: Third-order groups and their contributions for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, AiT, ω, and V_{mo}

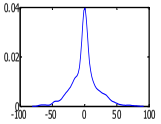
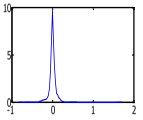
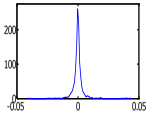
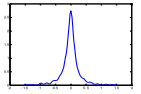
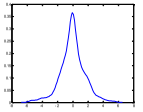
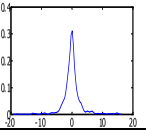
Table A8. CI method based property models: atom contributions and model constants for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, ω, and V_{mo}

Table A1. Performance of MG method based property models for thermo-physical and transport properties analysed using simultaneous regression method

Sl. No.	Property	L.H.S. of MG method based property prediction model	MG group-contribution model $f(X)=\sum_i N_i C_i+\sum_j M_j D_j+\sum_k E_k O_k$											
		$f(X)$	N	v	R^2	Residual distribution plot	P_{rc} (± 1%)	P_{rc} (± 5%)	P_{rc} (± 10%)	SD	AAE	ARE [a]	AE_{max}	RE_{max}
1	Normal boiling point [K]	$exp\left(\frac{T_b}{T_{bo}}\right)$	3510	3178	0.9989		59.38	99.47	100	7.67	5.96	1.38	33.47	11.39
2	Critical temperature [K]	$exp\left(\frac{T_c}{T_{co}}\right)$	858	607	0.9966		66.14	98.87	100.0	10.91	7.82	1.26	44.34	7.21
3	Critical pressure [bar]	$\left(P_c-P_{c1}\right)^{-0.5}-P_{c2}$	852	608	0.9808		38.06	81.19	92.91	1.85	1.16	3.39	11.79	29.82
4	Critical volume [cc/mol]	V_c-V_{c0}	797	565	0.9959		52.45	93.35	99.50	11.30	7.71	2.0	61.19	20.57

5	Normal melting point [K]	$exp\left(\frac{T_m}{T_{mo}}\right)$	5183	4803	0.9456		15.94	69.71	91.92	21.45	17.65	5.59	132.64	77.46
6	Gibbs free energy [kJ/mol]	$\Delta_f G^o - G_{f0}$	749	521	0.9987		42.46	78.91	87.98	7.59	4.57	**	39.69	**
7	Enthalpy of formation [kJ/mol]	$\Delta_f H^o{}_{gas} - H_{f0}$	882	649	0.9993		52.83	82.99	90.02	6.60	4.15	**	33.68	**
8	Enthalpy of fusion [kJ/mol]	$\Delta_{fus} H - H_{fus0}$	764	516	0.8477		13.61	35.60	54.97	4.92	2.58	**	69.09	**
9	Octanol/Water partition coefficient	$LogK_{ow} - K_{ow0}$	12193	11817	0.8825		5.68	25.64	44.82	0.61	0.45	**	4.18	**
10	Flash point [K]	$F_p - F_{po}$	512	340	0.9706		39.84	90.63	98.63	11.43	8.38	2.61	51.61	14.95

11	Hansen solubility parameter [MPa$^{1/2}$]	Dispersion (δ_D)	10.37	769	0.7374		42.53	88.52	96.82	0.94	0.53	**	9.59	**
		Polar(δ_P)	1017	754	0.7528		10.91	28.32	46.21	2.05	1.44	**	9.03	**
		H_2-bond (δ_H)	1016	754	0.8737	(shown here for δ_D parameter)	12.99	34.06	56.00	1.57	1.08	**	9.24	**
12	Enthalpy of vaporization (298K) [kJ/mol]	$\Delta_{vap}H^\circ - H_{v0}$	705	509	0.9751		45.82	92.20	97.16	2.19	1.16	2.91	27.10	134.83
13	Enthalpy of vaporization (Tb) [kJ/mol]	$\Delta_{vap}H - H_{vb0}$	512	346	0.9647		48.44	90.23	98.83	1.35	0.87	2.44	8.36	22.83
14	Entropy of vaporization (Tb) [J/mol K]	$\Delta_{vap}S - S_{vb0}$	512	346	0.8620		59.96	95.31	98.83	3.0	1.72	1.84	19.80	19.61
15	Hildebrand solubility parameter [MPa$^{1/2}$]	$\delta - \delta_0$	1384	1089	0.8352		34.61	72.54	92.41	1.60	1.05	5.44	11.22	29.09

16	Auto ignition temperature [b] [K]	T_{AiT}	570	385	0.9742		50.65	93.14	100.00	20.12	13.51	2.09	80.14	10.86
17	Acentric factor	$exp\left(\frac{\omega}{\omega_a}\right)^{\omega_b} - \omega_c$	1723	1422	0.8945		39.64	54.09	78.00	0.0991	0.054	11.35	1.67	76.99
18	Liquid molar volume [cc/kmol]	V_m - V_{m0}	1056	800	0.9971		76.23	93.37	98.77	0.0042	0.0021	1.84	0.0423	55.22
19	Liquid viscosity [mPa-s]	$\ln(\mu)$	522	339	0.98		21.37	41.52	62.1	1.00	0.37	25.50	1.45	29.6
20	Liquid surface tension [mN/m]	σ	546	373	0.95		41.2	83.54	93.64	1.28	0.92	3.35	4.23	21.9
21	Thermal conductivity [mW/mK]	k	336	198	0.90		55.63	80.6	94.1	9.7	5.9	4.81	10.63	19.25

[a] For $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, and $LogK_{ow}$, ARE is not reported since these properties contain both positive and negative values. For Hansen solubility parameters δ_D, δ_P, and δ_H, and liquid viscosity, ARE and REmax is not reported as these properties contain very small experimental values. [b] For auto ignition temperature, the right hand side of the MG model is different and it is given as,

$$T_{Ait} = Ait_1\, 10^{-\sum N_i C_{ai} + \sum M_j D_{aj} + \sum E_k O_{ak}} + Ait_2 + \sum N_i C_{bi} + \sum M_j D_{bj} + \sum E_k O_{bk}$$

Table A2. MG method based property models analysed using step-wise regression method: First-order groups and their contributions [a] for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, and $LogK_{ow}$

	Group	T_{bIi}	T_{cIi}	P_{cIi}	V_{cIi}	T_{mIi}	$\Delta_f G^o_{Ii}$	$\Delta_f H^o_{gas\ Ii}$	$\Delta_{fus}H_{Ii}$	
1	CH_3	0.8853	1.0893	0.0052	58.5875	0.6699	-14.4011	-57.2760	1.0523	0.3008
2	CH_2	0.5815	3.4607	0.0087	56.5948	0.2992	8.1227	-20.8404	2.6577	0.4352
3	CH	-0.0039	4.6670	0.0123	46.1415	-0.2943	25.9816	7.3487	0.4978	0.3837
4	C	-0.4985	6.6186	0.0150	36.5473	-0.0430	51.2378	38.4721	-0.5662	0.6325
5	CH_2=CH	1.4381	5.2032	0.0109	100.5126	1.0511	79.1413	42.9281	1.1842	0.5059
6	CH=CH	1.1867	8.2561	0.0155	97.8555	0.9974	93.1905	68.7360	5.0742	0.7652
7	CH_2=C	1.0223	7.3562	0.0138	93.5562	0.7252	85.5619	63.3152	2.5470	0.6909
8	CH=C	0.7960	10.0151	0.0197	93.1129	0.3942	107.0623	95.6578	3.9264	0.7270
9	C=C	0.4825	13.5343	0.0470	109.4707	0.2490	128.5800	125.2762	4.0203	0.7043
10	CH_2=C=CH	2.2128	11.0153	0.0123	133.1546	1.7798	213.5337	184.0134	6.6534	****
11	CH_2=C=C	1.7877	12.6592	0.0195	145.8207	1.6984	227.6405	208.4542	7.6250	****
12	C=C=C	****	****	****	****	****	****	****	****	****
13	CH≡C	1.5899	6.5967	-0.0033	74.2536	2.2487	214.8806	211.9192	-0.8684	-0.1819
14	C≡C	1.3528	11.6990	0.0050	74.1507	1.4582	216.8223	227.9484	6.3111	0.4980
15	aCH	0.7302	3.7648	0.0016	43.6099	0.6092	21.5796	7.7572	1.5883	0.2224
16	aC fused with aromatic ring	1.1169	17.9666	0.0061	27.3290	1.8684	24.4475	26.3845	0.9534	0.3495
17	aC fused with non-aromatic ring	0.7651	18.2578	0.0053	23.9198	1.4145	32.3486	24.0608	1.7831	0.3370
18	aC except as above	0.8589	21.0933	0.0099	24.4318	0.7650	31.5588	33.2646	-0.0756	0.3005
19	aN in aromatic ring	1.1087	9.5991	-0.0086	15.5252	2.0810	83.5681	57.2732	12.8371	-0.4661
20	aC-CH_3	1.2884	7.9735	0.0121	94.6685	1.0601	19.0529	-23.8708	2.5389	0.6065
21	aC-CH_2	0.9449	11.8635	0.0213	78.6325	-0.0147	41.7191	14.0530	2.1892	0.5298
22	aC-CH	0.2543	12.1461	0.0242	78.6091	-0.3962	56.9679	42.3770	1.4423	0.6835
23	aC-C	-0.2719	7.1936	0.0341	68.2761	-0.5727	87.2843	77.7732	0.4054	0.8470
24	aC-CH=CH_2	1.9250	14.8290	0.0189	116.8456	1.2457	107.6121	73.3194	4.7879	1.3057
25	aC-CH=CH	1.9128	22.0531	0.0280	106.7447	1.8394	120.6419	100.0113	10.1196	0.8924
26	aC-C=CH_2	1.4666	16.6811	0.0266	94.3026	1.3068	125.1344	101.6121	4.7055	0.9316
27	aC-C≡CH	1.8337	16.9735	0.0102	85.9776	1.8603	255.3446	253.3361	3.2779	0.9306
28	aC-C≡C	2.0843	****	****	****	5.0389	****	****	****	1.2057
29	OH	2.1385	10.1673	-0.0146	19.6860	3.2702	-161.2347	-192.3701	4.0062	-1.0185
30	aC-OH	2.5726	24.0563	-0.0125	42.4182	5.1868	-138.0599	-171.1375	8.5193	-0.0081
31	COOH	4.0158	30.6324	-0.0040	81.8277	7.1886	-353.9250	-406.6250	9.9005	-0.9393
32	aC-COOH	4.6109	43.9599	0.0088	102.5828	14.2024	-318.9761	-370.3788	21.1395	0.1389
33	CH_3CO	2.6245	14.1935	0.0072	120.0335	3.1357	-138.6683	-195.3595	7.0644	-0.3774
34	CH_2CO	2.0151	15.1212	0.0124	120.5818	2.9007	-119.9226	-162.8240	8.9695	-0.1855
35	CHCO	1.3147	14.7966	0.0254	107.4178	1.4295	-99.0503	-124.8225	8.2525	0.2045
36	CCO	0.7443	****	****	****	1.9060	****	-198.3977	****	0.1934
37	aC-CO	2.3558	29.9399	0.0156	67.3860	2.5839	-82.2708	-100.2815	3.8550	-0.0915
38	CHO	2.1663	11.2211	-0.0046	55.0894	2.9884	-116.7265	-143.3957	9.3980	-0.7850
39	aC-CHO	2.7098	25.4370	0.0110	102.5369	3.0709	-87.1171	-110.4720	7.8333	-0.1009
40	CH_3COO	2.5805	12.9825	0.0149	138.8535	2.4227	-322.9761	-402.3041	7.5327	-0.4615
41	CH_2COO	2.1808	15.6211	0.0175	132.7591	1.5439	-297.5792	-364.9735	9.5106	-0.2893

42	CHCOO	1.5283	15.3752	0.0224	134.6872	1.0130	-283.3789	-337.2026	10.1226	-0.3862
43	CCOO	0.6985	****	****	****	0.5572	****	****	****	0.1050
44	HCOO	2.2411	11.5499	-0.0028	82.1238	2.3697	-293.6109	-342.2988	7.7387	-0.7585
45	aC-COO	2.1906	20.7712	0.0205	95.8996	1.5433	-279.6779	-300.5819	3.8902	-0.0459
46	aC-OOCH	****	****	****	****	****	****	****	****	-0.3394
47	aC-OOC	2.1165	23.6575	0.0211	131.3026	3.6591	****	****	7.2555	-0.2187
48	COO except as above	1.6694	11.7021	0.0050	80.2600	1.6224	-298.0974	-331.8846	10.5508	-0.4766
49	CH_3O	1.5724	5.6588	0.0035	77.1882	1.5327	-108.4394	-171.5125	4.5423	-0.3030
50	CH_2O	0.9999	6.2957	0.0112	72.8224	0.7649	-103.5032	-150.9552	4.9397	-0.1449
51	CH-O	0.4724	5.4860	0.0502	77.4178	0.1817	-91.8503	-132.6225	8.1225	0.1650
52	C-O	-0.1490	7.4900	0.0455	70.7497	0.0556	-57.6156	-91.9940	4.4119	0.8028
53	aC-O	1.2119	13.0936	0.0104	33.6376	1.3409	-73.2701	-85.4935	-1.2987	-0.0128
54	CH_2NH_2	2.3212	12.9223	-0.0043	111.9231	3.4368	52.0484	-25.6857	11.9943	-1.4650
55	$CHNH_2$	1.5629	11.2234	0.0057	75.8207	37.8806	62.3492	-0.3708	6.4611	-1.7988
56	CNH_2	1.0473	11.0799	0.0198	89.1457	11.4243	73.4003	16.7802	****	-1.4858
57	CH_3NH	1.9861	10.6027	0.0030	92.8081	2.7205	75.7710	-4.8071	5.0200	-0.6853
58	CH_2NH	1.3838	9.7537	0.0154	103.0971	2.0673	89.1152	23.1822	6.3141	-0.9465
59	CHNH	0.7116	8.7789	0.0199	109.4178	1.6571	90.8497	36.5775	3.6025	-0.4419
60	CH_3N	1.0505	8.1028	0.0152	87.8679	0.9396	125.2589	52.5941	6.2150	-0.3519
61	CH_2N	0.4199	7.5227	0.0383	85.6666	-0.1982	137.6484	76.7801	2.7449	-0.6373
62	aC-NH_2	2.8950	27.9618	-0.0007	70.7212	4.3691	60.7058	12.9910	5.3925	-0.2684
63	aC-NH	1.9748	27.0506	0.0060	70.4515	2.2171	107.8738	76.4963	2.2418	0.3784
64	aC-N	1.2017	18.8267	0.0175	101.6276	1.8403	169.9541	123.9093	3.1432	0.0704
65	NH_2 except as above	1.8748	9.4911	-0.0202	47.6500	3.5553	22.9886	18.5210	6.4001	-0.7775
66	CH=N	1.1381	****	****	****	9.6493	****	****	****	0.1666
67	C=N	0.8619	****	****	****	1.5618	****	****	****	0.6829
68	CH_2CN	3.5887	26.0365	0.0048	136.0537	2.7022	120.9650	80.7816	5.9251	-0.4782
69	CHCN	2.8741	21.2156	0.0169	134.7729	1.8370	143.3574	127.6662	9.7480	-0.0320
70	CCN	2.0085	18.0027	0.0269	131.0479	2.6597	160.4040	138.7443	5.1387	-0.2908
71	aC-CN	3.1428	29.1970	0.0113	92.9776	5.1372	154.3146	162.0197	4.7079	-0.0376
72	CN except as above	2.6954	14.6317	0.0000	76.1410	3.4753	113.8291	114.0855	5.4938	-0.0185
73	CH_2NCO	2.9220	****	****	****	****	****	****	****	****
74	CHNCO	2.3740	****	****	****	****	****	****	****	****
75	CNCO	****	****	****	****	10.1488	****	****	****	****
76	aC-NCO	2.4141	17.1140	0.0146	127.5953	2.5263	9.4446	-22.5639	****	****
77	$CH2NO_2$	3.7481	24.3244	0.0009	147.1546	2.8903	****	-80.3757	10.5773	-0.1906
78	$CHNO_2$	2.9440	24.1232	0.0073	142.8207	0.8742	17.3405	-45.6885	****	-0.1587
79	CNO_2	2.2443	****	****	****	6.6444	****	-39.9148	-3.2591	-0.0837
80	aC-NO_2	3.5108	33.7161	0.0097	122.7916	4.5166	45.3163	-22.3281	7.2988	0.1414
81	NO_2 except as above	2.9569	24.3580	****	****	3.4486	****	-52.4431	5.5642	-0.3574
82	ONO	1.7563	****	****	****	****	****	****	****	****
83	ONO_2	2.7717	****	****	****	2.5055	****	-111.5866	9.1872	0.0183
84	$HCON(CH_2)_2$	4.5451	****	****	****	****	****	****	****	****
85	$HCONHCH_2$	5.9827	****	****	****	****	****	****	****	-1.5813
86	$CONH_2$	5.3258	64.8591	-0.0166	128.4957	12.1011	-143.8493	-210.9634	16.4373	-1.3025
87	$CONHCH_3$	4.2507	50.9596	0.0021	180.4957	6.4777	-119.2493	-217.9018	17.0273	-0.5040

88	$CONHCH_2$	5.1104	****	****	****	5.0152	****	-183.6249	****	-0.8272
89	$CON(CH_3)_2$	4.7928	36.3200	0.0176	234.4957	3.7968	-72.7493	-202.9018	11.1373	-0.3265
90	$CONCH_3CH_2$	****	****	****	****	2.1190	****	****	****	-0.7826
91	$CON(CH_2)_2$	3.4965	****	****	****	2.3854	****	****	****	-0.3486
92	CONHCO	5.8523	****	****	****	9.5767	****	****	****	-1.8638
93	CONCO	4.1284	****	****	****	3.6218	****	****	****	-1.0103
94	$aC\text{-}CONH_2$	6.3545	****	****	****	14.9334	****	-167.9009	17.5979	-1.0168
95	aC-NH(CO)H	5.7016	57.2710	0.0132	135.9776	6.4873	-52.4554	-129.1639	9.4379	-0.3782
96	aC-N(CO)H	3.8073	****	****	****	3.9043	****	****	****	-0.6462
97	aC-CONH	5.5102	****	****	****	7.9133	****	****	15.5818	-0.5608
98	aC-NHCO	5.4143	****	****	****	9.5142	****	****	5.6651	-0.2977
99	aC-(N)CO	****	****	****	****	6.8188	****	****	****	-1.1225
100	NHCONH	****	****	****	****	9.9235	****	****	10.0373	-1.2694
101	NH_2CONH	****	****	****	****	14.4374	****	****	14.2234	-1.2441
102	NH_2CON	****	****	****	****	11.9495	****	****	11.1196	-0.8436
103	NHCON	5.2529	****	****	****	8.0774	****	****	12.9026	-0.4984
104	NCON	2.7488	****	****	****	3.9673	****	-11.6438	10.9703	-1.4325
105	$aC\text{-}NHCONH_2$	4.4378	****	****	****	18.9947	****	****	17.4879	-0.6146
106	aC-NHCONH	0.7622	****	****	****	23.7375	****	****	20.5918	-0.0958
107	NHCO except as above	****	****	****	****	4.3405	****	****	****	-1.0853
108	CH_2Cl	2.2351	11.9702	0.0049	102.1071	1.9141	-39.7754	-84.7131	4.2198	0.6337
109	CHCl	1.5459	11.5936	0.0076	100.5343	1.2675	-29.6967	-62.7806	****	0.9663
110	CCl	0.9550	9.0702	0.0184	96.1457	1.5407	-19.6497	-48.1198	0.7126	0.6949
111	$CHCl_2$	2.7533	16.9240	0.0082	149.4565	2.5271	-45.1546	-96.6848	6.5593	0.9682
112	CCl_2	2.1881	****	****	****	3.3617	****	****	2.0050	0.8876
113	CCl_3	3.1119	18.9946	0.0128	194.4957	4.2995	-60.4593	-120.2018	3.0773	1.7802
114	CH_2F	1.5304	6.6907	-0.0015	77.4957	1.7066	-196.5493	-242.3018	6.7273	-0.0374
115	CHF	****	5.2654	-0.0371	****	-0.0079	****	****	****	-0.2019
116	CF	****	****	****	****	****	****	****	****	****
117	CHF_2	1.6770	6.2521	0.0065	80.6485	1.9419	-427.5493	-478.7018	6.5973	1.1936
118	CF_2	0.4099	2.3530	0.0089	93.6941	0.3710	****	****	****	0.7860
119	CF_3	1.1029	1.1504	0.0306	98.8850	1.1201	-624.4337	-684.3807	3.0069	0.8799
120	CCl_2F	2.1877	9.2822	0.0111	161.1293	3.3808	-267.2334	-316.2920	2.7953	1.6157
121	HCClF	1.8183	****	****	****	1.2161	****	****	****	0.5822
122	$CClF_2$	1.5105	5.3326	0.0238	135.9554	1.8240	-412.2168	-478.1757	1.4443	1.1337
123	aC-Cl	1.6978	16.8718	0.0081	79.3282	1.9748	-6.5257	-17.7872	4.3158	0.8999
124	aC-F	0.6956	2.8684	0.0066	50.4045	0.9178	-175.5854	-165.1963	1.9020	0.3494
125	aC-I	2.9857	34.1350	0.0061	104.9776	2.4680	83.6446	90.9361	3.5779	0.7981
126	aC-Br	2.2711	17.4679	0.0028	79.9336	2.5918	35.4709	36.0808	6.4040	0.9655
127	-I except as above	2.6464	17.5564	-0.0081	94.0868	2.4184	38.0332	43.6127	5.5859	0.7383
128	-Br except as above	2.0032	11.7189	-0.0154	67.8250	1.9506	-10.4356	-16.0308	3.6841	0.4456
129	-F except as above	0.7259	-0.3201	-0.0023	15.5722	0.9410	-199.9915	-217.0159	4.9754	0.0816
130	-Cl except as above	1.3550	5.2618	-0.0073	46.2125	1.4383	-46.9553	-63.0636	4.6208	0.3245
131	CHNOH	3.7873	****	****	****	4.1895	****	****	****	-0.1279
132	CNOH	3.1862	****	****	****	4.4116	****	****	****	-0.5409
133	aC-CHNOH	****	****	****	****	9.8629	****	****	****	-0.0934

134	OCH_2CH_2OH	3.8550	19.6040	0.0083	149.0103	2.4711	-251.5191	-361.0753	7.1850	-1.0987
135	$OCHCH_2OH$	3.1854	****	****	****	10.7229	****	****	****	-0.7924
136	OCH_2CHOH	2.8513	17.9277	0.0136	148.1113	1.9288	-238.7942	-333.6449	12.9941	-0.6665
137	-O-OH	3.0402	10.9864	-0.0132	49.7748	4.1185	****	-144.8219	****	****
138	CH_2SH	2.6797	16.1811	0.0008	98.0676	2.1968	11.2800	-22.7338	8.7841	0.4238
139	CHSH	1.9567	15.0366	0.0022	108.8207	0.8159	27.9605	3.4742	4.4411	-0.0033
140	CSH	1.3166	15.2244	0.0146	103.1457	5.3745	45.5303	27.3502	1.1026	-0.1311
141	aC-SH	2.5609	25.5454	0.0027	88.9628	2.8310	41.0446	37.5361	5.2879	0.5517
142	-SH except as above	2.0936	13.1916	-0.0166	51.4055	2.4561	-3.8764	-3.4198	5.9117	0.3627
143	CH_3S	2.5139	14.4857	0.0001	111.1688	1.9363	15.6214	-18.1453	7.1965	0.5282
144	CH_2S	2.0027	16.7767	0.0099	106.2861	0.9493	34.8651	11.6152	4.2728	0.4409
145	CHS	1.3865	****	****	****	0.9275	****	****	****	0.7805
146	CS	0.8100	****	****	****	1.7141	****	****	****	1.3864
147	aC-S-	1.8625	28.7817	0.0113	****	1.1686	****	****	****	0.6057
148	SO	4.8492	53.1191	-0.0106	81.8207	5.5058	-51.2995	-71.0858	13.6050	-1.0179
149	SO2	5.1615	48.1622	0.0002	87.8483	6.7916	-256.9727	-305.4737	17.9450	-0.6102
150	SO_3 (sulfite)	2.9700	21.9243	0.0030	114.6314	****	****	-430.8449	****	0.3189
151	SO_3 (Sulfonate)	5.1010	****	****	****	6.1276	****	****	****	-1.2718
152	SO_4 (Sulfate)	4.2212	50.7390	-0.0296	143.7272	3.9937	-518.9821	-621.4354	****	-0.8196
153	aC-SO	4.4416	****	****	****	5.2165	****	****	****	-0.6632
154	$aC\text{-}SO_2$	****	****	****	120.2048	4.5946	-305.3319	-360.5938	4.9168	-0.5006
155	PH (phosphine)	****	****	****	****	****	****	****	****	****
156	P (Phospine)	0.8503	****	****	****	0.3443	****	****	****	****
157	PO_3 (Phospite)	1.9015	****	****	****	****	****	****	****	-1.7413
158	PHO_3 (Phosponate)	4.3669	****	****	****	0.1112	****	****	****	-2.0081
159	PO_3 (Phosponate)	2.9750	****	****	****	0.5547	****	****	****	-1.0133
160	PHO_4 (Phospate)	****	****	****	****	3.0670	****	****	****	-2.0272
161	PO_4 (Phospate)	2.6270	49.3123	****	94.2999	2.0934	****	-1045.5285	****	-1.7943
162	$aC\text{-}PO_4$	****	****	****	****	-2.1633	****	-975.3652	7.7157	-0.8714
163	aC-P	****	158.1164	****	-177.160	1.0341	****	102.1348	-2.2043	0.8175
164	CO_3 (Carbonate)	2.2704	16.9445	0.0066	101.7272	2.8869	-436.5071	-518.0449	11.6750	-1.0785
165	C_2H_3O	2.5872	13.5396	0.0024	114.9011	1.4240	-5.2419	-67.0613	****	-0.4950
166	C_2H_2O	1.9387	15.4199	0.0037	193.8207	-2.1100	2.8205	-51.4258	****	****
167	C_2HO	****	****	****	****	-0.8025	****	****	****	-0.2313
168	CH_2 (cyclic)	0.7067	3.6309	0.0037	46.4028	0.5067	9.2444	-22.6402	1.0024	0.1818
169	CH (cyclic)	0.3922	4.7667	0.0113	39.9190	0.2691	14.1714	-6.0339	1.1929	0.2934
170	C (cyclic)	-0.2034	3.6897	0.0137	47.2221	0.5775	22.2696	24.9886	-0.3560	0.2412
171	CH=CH (cyclic)	1.3880	6.4670	0.0023	78.3373	1.2370	77.5968	55.7156	1.0288	0.3902
172	CH=C (cyclic)	0.9412	9.1584	0.0079	74.6735	0.7830	62.0945	41.5909	4.0736	0.5383
173	C=C (cyclic)	0.6681	****	****	****	0.8144	****	****	****	0.6519
174	CH_2=C (cyclic)	1.2991	10.0648	0.0034	89.6110	0.9182	****	****	4.5224	-0.2412
175	NH (cyclic)	1.7300	9.1566	-0.0087	45.8306	3.8454	71.2591	36.8801	6.8973	-0.5426
176	N (cyclic)	1.0768	7.2834	0.0008	33.4400	0.6005	85.8513	72.9845	-1.1760	-0.6373
177	CH=N (cyclic)	5.5183	24.3803	0.0259	****	5.0372	****	58.1685	3.9592	-0.1217
178	C=N (cyclic)	4.9471	****	****	****	6.3416	****	59.6723	****	0.1050
179	O (cyclic)	0.8691	6.3363	-0.0061	13.9679	1.3269	-121.1936	-132.1780	3.2867	-0.4008

180	CO (cyclic)	2.2409	33.8840	-0.0070	51.3884	3.4512	-165.5643	-198.1806	5.6624	-0.2971
181	S (cyclic)	1.8655	11.8984	-0.0066	33.5620	1.6341	3.3634	15.9851	4.7493	0.3935
182	SO_2 (cyclic)	****	68.0111	-0.0077	87.9154	5.9517	-258.2402	-301.2129	9.2060	-1.0227
183	>NH	****	****	****	****	1.5837	****	****	****	-0.5436
184	-O-	1.1466	10.6446	0.0002	12.6373	0.6385	-107.8952	-121.5559	1.7178	-0.2784
185	-S-	1.2669	****	****	****	2.4091	****	****	****	0.7268
186	>CO	1.1503	****	****	****	1.6842	****	****	****	-0.0433
187	PO_2	1.4250	****	****	****	****	****	****	****	****
188	CH-N	0.1249	****	****	****	-0.5463	****	****	****	-0.4053
189	SiHO	0.1000	****	****	****	-0.7266	****	****	****	****
190	SiO	-0.2162	****	****	****	-1.2440	****	****	****	0.0526
191	SiH_2	1.0050	****	****	****	0.4337	****	****	****	****
192	SiH	0.3938	****	****	****	0.4074	****	****	****	0.9170
193	Si	-0.3351	7.8515	0.0366	****	0.9350	****	****	****	1.5213
194	$(CH_3)_3N$	****	****	****	****	****	****	****	****	****
195	N=N	****	****	****	****	-2.0114	****	194.8624	****	-0.4148
196	C_{cyclic}=N-	1.4245	****	****	****	1.2769	****	197.5934	****	0.0819
197	C_{cyclic}=CH-	1.0676	****	****	****	0.3162	****	****	****	0.0782
198	C_{cyclic}=NH	****	****	****	****	****	****	****	****	-0.5150
199	N=O	3.1083	****	****	****	1.5761	****	****	****	-0.1673
200	C_{cyclic}=C	0.4528	17.0872	0.0348	****	0.1669	****	****	****	0.5066
201	P=O	****	****	****	****	4.2703	****	****	****	-0.4081
202	N=N	****	****	****	****	0.3581	****	249.6976	****	0.7929
203	C=NH	****	****	****	****	4.2948	****	****	****	-0.0797
204	>C=S	****	****	****	****	2.6644	****	94.6074	****	0.4673
205	aC-CON	****	****	****	****	9.1484	****	****	****	-1.3222
206	aC=O	****	****	****	****	5.0141	****	****	10.4711	-0.2678
207	aN-	****	****	****	****	2.6021	****	211.9535	****	-2.1639
208	-Na	****	****	****	****	****	****	****	****	****
209	-K	****	****	****	****	****	****	****	****	****
210	HCONH	****	****	****	****	9.9337	****	****	****	-0.9190
211	CHOCH	****	****	****	****	0.9204	****	****	****	****
212	C_2O	****	****	****	****	****	****	****	****	****
213	SiH_3	****	****	****	****	****	****	****	****	****
214	SiH_2O	****	****	****	****	****	****	****	****	****
215	CH=C=CH	****	****	****	****	****	****	****	****	****
216	CH=C=C	****	****	****	****	****	****	****	****	****
217	OP(=S)O	****	****	****	****	****	****	****	****	****
218	R	****	****	****	****	****	****	****	****	****
219	$CF_{2cyclic}$	0.5028	****	0.0084	****	0.9609	****	****	****	0.6839
220	CF_{cyclic}	0.6977	****	****	****	-0.8497	****	****	****	0.0743

[a] The symbols T_{b1i}, T_{c1i}, P_{c1i}, V_{c1i}, T_{m1i}, $\Delta_f G^o{}_{1i}$, $\Delta_f H^o{}_{gas1i}$, $\Delta_{fus} H_{1i}$, and $LogK_{ow1i}$ represent the contributions (C_i) of the first-order groups for the corresponding properties

Table A2 (Continued). MG method based property models analysed using step-wise regression method: First-order groups and their contributions [a] for the properties÷ F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, ω and V_m

	Group	F_{pIi}	$\Delta_{vap}H^o_{Ii}$	$\Delta_{vap}H_{Ii}$	$\Delta_{vap}S_{Ii}$	δ_{DIi}	δ_{PIi}	δ_{HIi}	δ_{Ii}	ω_{Ii}	V_{mIi}
1	CH_3	21.7458	2.2643	2.0701	-0.1822	7.5697	1.9996	2.2105	-2.1040	0.0077	0.0241
2	CH_2	11.5194	4.7607	2.3353	0.6185	-0.0018	-0.1492	-0.2150	-0.1540	0.0047	0.0165
3	CH	-5.1205	5.0336	1.6963	0.2664	-7.7208	-2.7099	-2.6826	1.1153	0.0005	0.0086
4	C	-19.7535	5.1448	0.8251	0.0320	-15.4498	-4.7191	-6.4821	1.9577	-0.0069	0.0007
5	CH_2=CH	32.1079	5.4644	4.1230	0.8081	7.6415	2.4926	2.7673	-2.6837	0.0098	0.0347
6	CH=CH	20.4241	8.1390	4.1052	1.7154	-0.0393	-0.1344	0.8631	-0.7619	0.0101	0.0269
7	CH_2=C	11.6312	7.3334	3.3750	1.0404	-0.1038	0.0444	1.0623	-1.3323	0.0026	0.0264
8	CH=C	0.2590	8.7141	3.9363	0.8931	-6.9792	-1.2242	-0.1204	1.2028	0.0016	0.0161
9	C=C	-9.8452	10.7625	4.0746	1.3878	-14.4343	-6.5194	-2.9995	1.9713	-0.0112	0.0132
10	CH_2=C=CH	****	11.3320	6.5314	1.4470	7.8035	2.4810	3.6733	-3.3243	0.0130	0.0428
11	CH_2=C=C	****	****	****	****	-0.0395	-1.4993	0.0791	-1.5972	0.0054	0.0359
12	C=C=C	****	****	****	****	****	****	****	****	****	****
13	CH≡C	50.2118	6.6654	5.6796	3.4692	7.8142	3.1014	3.7484	-1.8143	0.0121	0.0263
14	C≡C	****	11.0141	7.8319	4.7793	-0.2242	-0.3012	1.9582	0.8547	0.0079	0.0153
15	aCH	17.6554	4.1504	2.6770	0.5379	3.1144	0.9945	0.7694	-0.2272	0.0041	0.0122
16	aC fused with aromatic ring	19.3539	6.4425	4.8655	2.7828	-2.5088	-2.3684	-2.7569	0.2066	0.0027	0.0066
17	aC fused with non-aromatic ring	3.1895	5.7028	3.3349	1.3254	-2.2634	-1.6748	-2.6898	-0.3397	0.0021	0.0060
18	aC except as above	17.8540	6.9303	****	****	-7.0934	-3.8082	-1.4992	-0.3453	0.0033	0.0079
19	aN in aromatic ring	43.4662	10.7328	6.7992	6.4441	3.3472	5.2498	3.9266	-0.4729	0.0037	0.0050
20	aC-CH_3	30.8807	8.1461	4.4356	-0.3329	3.0551	-0.1775	0.1713	-0.7833	0.0094	0.0282
21	aC-CH_2	23.3343	9.9269	2.8182	-2.0971	-5.3239	-3.5919	-3.1805	0.4007	0.0083	0.0214
22	aC-CH	6.4828	9.2823	****	****	-12.9139	-5.1314	-7.0890	1.0885	0.0005	0.0144
23	aC-C	4.5266	12.6547	4.6514	1.7047	-19.5489	-7.5530	-7.8370	2.7336	-0.0067	0.0080
24	aC-CH=CH_2	48.4825	****	9.8956	6.4952	2.9119	-1.4985	1.0566	-1.0001	0.0129	0.0376
25	aC-CH=CH	48.1746	****	****	****	-4.5593	-3.1909	-3.6608	0.2022	0.0106	0.0287
26	aC-C=CH_2	31.6544	****	****	****	****	****	****	-0.2763	0.0085	0.0301
27	aC-C≡CH	45.9443	****	****	****	3.2278	-2.1724	0.1531	-1.7236	0.0049	0.0336
28	aC-C≡C	****	****	****	****	-3.9509	****	-2.9946	-0.8352	0.0000	0.0413
29	OH	78.5878	24.1639	16.4887	19.4039	8.0236	4.9598	11.8005	2.4637	0.0502	0.0044
30	aC-OH	83.3745	37.9847	15.7442	13.4747	3.1695	2.2263	7.7635	1.9922	0.0313	0.0142
31	COOH	115.3016	20.4196	****	****	8.4676	3.5123	7.5917	3.1694	0.0553	0.0244
32	aC-COOH	148.9109	****	****	****	3.0454	1.6294	4.5546	1.6303	0.0566	0.0468
33	CH_3CO	70.9382	15.1303	9.6190	2.6552	8.1630	6.0520	3.4394	-0.1581	0.0220	0.0345
34	CH_2CO	67.4790	16.4955	10.1585	4.7828	0.5557	0.7632	-0.0788	0.5639	0.0204	0.0288
35	CHCO	****	12.4208	9.1433	4.4806	****	****	****	2.4461	0.0164	0.0235
36	CCO	****	13.7977	1.4869	-8.0628	-15.3821	****	****	1.6430	-0.0032	0.0247
37	aC-CO	61.1990	22.7719	****	****	-4.0748	0.7990	-0.8306	1.7323	0.0195	0.0187
38	CHO	62.3286	13.2496	9.3020	4.6854	7.7699	7.9891	5.3761	0.9642	0.0228	0.0169
39	aC-CHO	76.9443	****	****	****	3.7310	4.4838	2.3964	1.2492	0.0270	0.0162
40	CH_3COO	73.7009	20.0986	13.0656	9.0305	8.0220	2.8480	5.0132	-0.6905	0.0295	0.0412
41	CH_2COO	50.2088	20.8219	13.5635	9.6708	0.4586	1.4477	2.7824	1.1280	0.0284	0.0365

42	CHCOO	30.2375	21.7450	11.1043	6.7688	-7.3196	-1.8452	-1.1517	1.7246	0.0183	0.0291
43	CCOO	****	18.7725	9.0957	5.8422	-15.1450	-7.8478	-10.5735	****	0.0135	0.0246
44	HCOO	57.4720	15.7192	10.8115	8.9224	7.9333	5.1029	6.8448	-0.4475	0.0242	0.0259
45	aC-COO	45.3229	22.9419	15.6863	16.7877	-5.1676	0.5739	-0.5041	1.5495	0.0170	0.0287
46	aC-OOCH	****	****	****	****	****	****	****	****	****	****
47	aC-OOC	****	****	****	****	-3.7074	-1.3895	-0.0219	2.1020	0.0199	0.0249
48	COO except as above	53.3130	17.1781	11.2832	8.7093	0.2832	1.6464	2.6377	1.8624	0.0256	0.0199
49	CH_3O	41.9635	7.8029	5.6432	1.9768	7.6577	3.0860	3.3464	-1.7006	0.0173	0.0283
50	CH_2O	32.9140	8.5931	5.1006	2.7380	0.1978	0.6423	0.8246	-0.0286	0.0160	0.0228
51	CH-O	-8.9309	9.1274	4.3695	3.3277	-7.7099	-1.9180	-2.1543	0.5182	0.0026	0.0207
52	C-O	****	9.0063	3.5593	2.1936	****	****	****	2.7118	-0.0002	0.0099
53	aC-O	3.0689	13.9610	8.0941	5.4927	-4.5924	-0.5729	0.0149	1.9263	0.0154	0.0108
54	CH_2NH_2	63.0277	16.3428	10.6011	6.5612	8.1717	5.2964	6.7984	0.3765	0.0286	0.0281
55	$CHNH_2$	20.3665	13.9985	8.1472	6.8940	-0.3287	0.6603	2.8953	-0.3947	0.0192	0.0207
56	CNH_2	****	13.2358	9.4272	9.3683	****	****	****	0.3815	0.0085	0.0179
57	CH_3NH	52.6666	13.9706	9.5323	7.2790	8.1301	3.4132	7.2551	-0.3527	0.0206	0.0282
58	CH_2NH	38.5602	11.9165	7.2608	2.6173	0.2374	0.1072	1.4183	0.9615	0.0250	0.0260
59	CHNH	12.2191	10.9079	5.0019	2.2862	-7.7581	-3.5886	-2.2824	0.6915	0.0136	0.0209
60	CH_3N	49.4137	13.6543	4.4346	2.8790	0.4260	-0.5166	2.4585	-0.5351	0.0113	0.0259
61	CH_2N	-2.1262	7.4283	5.6474	4.9009	-6.6418	0.0847	-7.3014	-1.8019	0.0077	0.0187
62	aC-NH_2	99.5136	23.9175	13.3649	7.4738	4.1237	2.8781	6.3616	1.8406	0.0254	0.0232
63	aC-NH	54.4905	16.7337	****	****	-3.7891	-2.5983	3.2590	0.5370	0.0200	0.0102
64	aC-N	31.8530	17.9388	****	****	****	****	****	1.1306	0.0084	0.0190
65	NH_2 except as above	65.5706	17.5679	9.5708	8.1026	8.5025	5.5141	9.0778	3.7625	0.0206	0.0104
66	CH=N	****	****	****	****	****	****	****	****	****	****
67	C=N	****	****	****	****	-8.0972	-2.4853	-3.1673	-3.7310	****	****
68	CH_2CN	87.3868	22.2337	15.6253	3.8719	8.6869	9.9443	3.3144	2.7451	0.0336	0.0309
69	CHCN	****	22.9889	12.8298	2.9701	1.3373	4.8107	1.7584	5.0723	0.0327	0.0267
70	CCN	53.0560	20.9458	10.7683	2.6833	-7.0171	0.3807	0.4757	5.8088	0.0265	0.0232
71	aC-CN	****	****	17.0956	12.9437	3.4153	5.8218	0.6067	1.2666	0.0239	0.0272
72	CN except as above	79.3428	19.1331	12.5795	4.2788	8.5463	9.8892	1.9219	1.1595	0.0256	0.0161
73	CH_2NCO	****	****	****	****	7.8303	10.0004	0.2895	****	0.0334	0.0404
74	CHNCO	****	****	****	****	****	****	****	****	****	****
75	CNCO	****	****	****	****	****	****	****	****	****	****
76	aC-NCO	****	****	****	****	3.4759	2.1596	1.8845	0.2311	0.0248	0.0311
77	$CH2NO_2$	107.8590	29.7120	19.5930	13.1928	8.9684	11.3856	3.1398	0.3261	0.0359	0.0327
78	$CHNO_2$	83.2334	27.1989	17.2398	10.6103	0.9349	6.5029	0.8294	3.0195	0.0289	0.0263
79	CNO_2	****	-38.8406	-24.2443	-4.1202	-6.6230	-0.9748	-0.5293	-3.4905	-0.0048	0.0513
80	aC-NO_2	105.7265	24.6451	****	****	3.6874	3.2054	1.2032	0.9670	0.0323	0.0278
81	NO_2 except as above	117.5232	26.3865	16.5196	7.6247	7.4714	4.4196	2.9346	0.1278	0.0297	0.0177
82	ONO	****	****	****	****	****	****	****	****	****	****
83	ONO_2	****	****	****	****	8.0429	7.3747	2.7871	-0.2537	0.0435	0.0280
84	$HCON(CH_2)_2$	****	36.1789	****	****	0.8140	6.4490	1.6091	4.2425	****	****
85	$HCONHCH_2$	****	46.5620	****	****	9.6303	8.0004	11.7895	8.8386	****	****
86	$CONH_2$	****	****	****	****	9.0445	9.3434	10.6398	10.8154	0.0427	0.0142
87	$CONHCH_3$	****	48.0983	****	****	7.4303	2.3004	3.4895	7.6437	0.0429	0.0380

88	$CONHCH_2$	****	50.9781	****	****	****	****	****	7.7425	****	****
89	$CON(CH_3)_2$	145.5232	38.3620	****	****	9.0320	9.1995	6.8046	2.7913	0.0342	0.0529
90	$CONCH_3CH_2$	****	****	****	****	****	****	****	****	****	****
91	$CON(CH_2)_2$	****	37.7058	****	****	-6.3092	5.3011	0.8686	4.9464	****	****
92	CONHCO	****	****	****	****	****	****	****	****	****	****
93	CONCO	****	31.1397	****	****	****	****	****	****	****	****
94	aC-$CONH_2$	****	****	****	****	5.6278	9.7276	7.3531	****	0.0453	0.0349
95	aC-NH(CO)H	****	****	****	****	****	****	****	4.1836	0.0436	0.0353
96	aC-N(CO)H	****	****	****	****	****	****	****	****	****	****
97	aC-CONH	****	****	****	****	****	****	****	****	****	****
98	aC-NHCO	80.7510	****	****	****	-3.9694	4.3121	3.8028	1.7478	0.0424	0.0317
99	aC-(N)CO	****	****	****	****	****	****	****	2.1061	****	****
100	NHCONH	****	****	****	****	****	****	****	****	****	****
101	NH_2CONH	****	****	****	****	10.9555	5.2566	8.1602	****	****	****
102	NH_2CON	****	****	****	****	****	****	****	****	****	****
103	NHCON	****	****	****	****	****	****	****	****	****	****
104	NCON	****	****	****	****	-13.5789	0.2015	2.1581	8.4504	****	****
105	aC-$NHCONH_2$	****	****	****	****	****	****	****	****	****	****
106	aC-NHCONH	****	****	****	****	****	****	****	****	****	****
107	NHCO except as above	****	44.9310	****	****	****	****	****	****	****	****
108	CH_2Cl	51.9241	13.1138	8.9579	3.4244	8.8860	4.1421	2.1489	-0.4616	0.0139	0.0317
109	CHCl	32.9439	10.4765	6.2408	0.9302	0.6503	3.0607	-0.5733	-0.2563	0.0077	0.0262
110	CCl	****	7.9858	3.3950	0.6453	-7.1092	1.6011	-6.9931	1.0190	-0.0018	0.0224
111	$CHCl_2$	80.3590	17.7306	11.2353	3.7878	9.2038	3.1389	2.5206	-0.8836	0.0141	0.0449
112	CCl_2	****	18.4989	****	****	1.1264	1.5615	-1.4311	0.4183	****	****
113	CCl_3	****	18.4766	12.2142	2.7297	9.8110	2.3695	0.1333	-2.4383	0.0129	0.0620
114	CH_2F	****	9.2076	8.9293	10.4628	7.2585	3.9074	-1.2477	-1.2343	0.0161	0.0249
115	CHF	****	****	****	****	-0.0364	-0.8934	-0.2549	****	-0.0055	0.0193
116	CF	****	****	****	****	****	****	****	****	****	****
117	CHF_2	****	10.0607	****	****	7.2270	7.1946	1.0771	-3.1844	0.0162	0.0345
118	CF_2	****	3.9712	2.3811	0.7881	-0.7620	****	****	-0.6044	0.0018	0.0252
119	CF_3	****	2.3578	0.8118	-0.4777	7.9050	3.2968	2.0149	-4.0611	0.0262	0.0413
120	CCl_2F	****	8.7800	4.8450	-1.4395	6.8167	-0.5158	****	-3.6258	0.0137	0.0564
121	HCClF	****	11.5260	7.0528	2.9386	7.8430	-0.1931	3.8252	-3.1021	0.0128	0.0391
122	$CClF_2$	****	9.7431	6.8539	2.3623	6.6167	1.1432	****	-4.1164	0.0145	0.0546
123	aC-Cl	52.7843	10.9174	6.5214	1.0767	3.6434	1.0768	0.6238	-0.2257	0.0090	0.0230
124	aC-F	****	4.4259	2.7810	1.1239	2.9601	0.8378	0.2767	-0.8265	0.0076	0.0167
125	aC-I	****	****	****	****	3.9278	1.0276	2.2531	-0.0830	0.0073	0.0351
126	aC-Br	66.0553	14.1751	****	****	3.7152	1.7417	1.6744	-0.0212	0.0087	0.0271
127	-I except as above	****	14.5775	9.6749	-3.9047	9.3425	2.6196	3.7401	0.0304	0.0083	0.0247
128	-Br except as above	61.7565	10.1516	6.4258	1.1044	9.0659	3.4215	2.9788	-0.8282	0.0094	0.0213
129	-F except as above	****	-2.5062	-0.2776	-0.2409	7.2762	2.7111	5.1354	-1.9563	0.0082	0.0138
130	-Cl except as above	37.4190	5.6012	4.2620	1.2260	7.9570	3.2931	1.8748	-1.1759	0.0075	0.0187
131	CHNOH	****	****	****	****	8.7303	2.0004	17.9895	****	0.0681	0.0224
132	CNOH	****	****	****	****	0.3614	0.3753	5.0366	****	****	****
133	aC-CHNOH	****	****	****	****	****	****	****	****	****	****

134	OCH_2CH_2OH	121.3382	31.3924	21.5721	14.9640	8.3416	5.8162	9.9562	1.8292	0.0560	0.0416
135	$OCHCH_2OH$	****	****	****	****	0.1491	0.0056	5.1167	****	0.0710	0.0371
136	OCH_2CHOH	****	****	****	****	0.2380	0.2947	4.8512	3.6034	0.0567	0.0394
137	-O-OH	15.8659	****	****	****	7.4303	13.0004	27.7895	4.3045	0.0693	0.0103
138	CH_2SH	56.8546	16.9958	10.4441	3.0567	8.6097	3.9160	4.2415	-0.3077	0.0135	0.0342
139	CHSH	****	15.1985	8.5222	2.3032	1.1605	2.8007	2.2675	-0.4829	0.0074	0.0293
140	CSH	****	14.4397	7.1115	1.2199	****	****	****	0.4329	0.0005	0.0252
141	aC-SH	****	17.1951	11.1256	4.2780	4.3851	0.7873	4.5741	0.6338	0.0092	0.0258
142	-SH except as above	****	11.7510	8.1071	2.5079	8.6744	2.5656	6.0326	-0.1659	0.0089	0.0188
143	CH_3S	48.3732	14.5642	9.5342	2.6917	8.7718	3.8085	3.2506	-1.0766	0.0129	0.0356
144	CH_2S	****	15.3044	9.1391	2.8468	0.7614	0.6080	0.1251	0.7398	0.0113	0.0295
145	CHS	****	14.4431	7.4714	3.4418	****	****	****	****	****	0.0690
146	CS	****	14.8261	4.6613	0.0112	****	****	****	****	-0.0009	0.0172
147	aC-S-	****	****	****	****	-3.5419	-2.1720	-1.3574	****	0.0118	0.0317
148	SO	147.0834	****	****	****	1.8357	7.6914	3.7447	9.2912	0.0056	0.0134
149	SO2	****	****	****	****	1.8166	8.4926	-0.3325	4.0124	0.0173	0.0228
150	SO_3 (sulfite)	****	****	****	****	****	****	****	1.8058	0.0345	0.0309
151	SO_3 (Sulfonate)	****	****	****	****	0.8820	2.7995	2.8046	****	-0.0128	0.0552
152	SO_4 (Sulfate)	141.0484	****	****	****	1.5623	11.9999	4.1941	3.8421	0.0182	0.0337
153	aC-SO	****	****	****	****	-3.6322	7.6074	-1.0695	****	****	****
154	aC-SO_2	****	****	****	****	-3.7817	7.9469	-0.3191	-10.7331	****	****
155	PH (phosphine)	****	****	****	****	****	****	****	****	****	****
156	P (Phospine)	****	****	****	****	-5.4711	-6.2794	****	****	0.0039	0.0192
157	PO_3 (Phospite)	****	****	****	****	****	****	****	1.7790	****	****
158	PHO_3 (Phosponate)	****	****	****	****	****	****	****	8.1425	****	****
159	PO_3 (Phosponate)	****	****	****	****	-6.0081	6.2614	0.6291	****	****	****
160	PHO_4 (Phospate)	****	****	****	****	****	****	****	****	****	****
161	PO_4 (Phospate)	100.1890	****	****	****	-6.2197	5.1066	2.1176	3.6009	0.0226	0.0306
162	aC-PO_4	-15.3995	****	****	****	-12.8906	3.8072	-1.7608	0.0443	0.0387	0.0560
163	aC-P	****	****	****	****	****	****	****	1.1108	-0.0156	0.0297
164	CO_3 (Carbonate)	****	19.9374	****	****	0.1623	1.2499	2.3941	1.9984	0.0305	0.0225
165	C_2H_3O	66.0416	16.0120	10.1680	6.0938	9.2764	5.6342	7.9801	1.1546	0.0204	0.0305
166	C_2H_2O	44.2334	****	****	****	0.9605	0.8007	2.0651	-0.4611	0.0034	0.0290
167	C_2HO	****	****	****	****	****	****	****	****	****	****
168	CH_2 (cyclic)	15.0958	3.8456	2.2775	-0.0801	2.6915	0.5026	0.6159	-0.3627	0.0036	0.0159
169	CH (cyclic)	10.5355	4.3257	2.4151	1.0078	-3.7719	-1.7549	-0.5171	-0.6398	0.0019	0.0063
170	C (cyclic)	-15.1444	3.5718	2.9031	3.2681	-7.1870	-2.2674	-2.6329	2.3564	-0.0089	0.0006
171	CH=CH (cyclic)	24.8992	7.7284	4.9731	1.4534	7.1887	1.2946	1.9775	-1.0024	0.0079	0.0253
172	CH=C (cyclic)	35.7375	6.6785	6.3925	5.2151	-0.8694	0.3059	-1.5025	1.0010	0.0074	0.0167
173	C=C (cyclic)	****	****	****	****	-9.3456	-11.5136	-17.0120	1.1243	0.0084	0.0141
174	CH_2=C (cyclic)	34.6184	5.5682	-6.2213	-22.7853	8.3422	4.1288	2.4163	-0.6471	0.0124	0.0321
175	NH (cyclic)	60.9244	16.3524	9.8928	8.0764	4.7659	3.0159	3.4090	1.7237	0.0145	0.0041
176	N (cyclic)	49.4820	****	****	****	-0.6101	1.2099	-0.6760	-0.1619	0.0046	0.0006
177	CH=N (cyclic)	****	12.0459	5.6900	2.0179	7.5861	8.8362	6.2565	-3.2969	0.0091	0.0214
178	C=N (cyclic)	****	****	****	****	****	****	****	****	0.0023	****
179	O (cyclic)	24.3976	6.6215	4.7567	3.9031	3.9616	3.1902	2.8020	0.5806	0.0111	0.0018

180	CO (cyclic)	76.3797	17.2410	20.5688	27.8334	2.7841	5.6981	2.7658	2.6466	0.0232	0.0130
181	S (cyclic)	****	13.2479	8.0085	2.1078	5.4675	2.6306	2.9683	0.8899	0.0062	0.0126
182	SO_2 (cyclic)	****	****	****	****	7.2362	11.9242	4.9204	5.0404	0.0245	0.0189
183	>NH	****	****	****	****	****	****	****	1.6621	0.0214	0.0104
184	-O-	-8.7112	****	****	****	-0.6898	-0.0731	0.2654	-7.4283	0.0019	0.0098
185	-S-	****	****	****	****	0.2257	-0.1052	-0.1864	****	0.0102	0.0157
186	>CO	****	20.0518	****	****	0.2109	0.5495	1.6226	-1.3882	0.0185	0.0122
187	PO_2	****	****	****	****	****	****	****	****	****	****
188	CH-N	****	****	****	****	****	****	****	****	****	****
189	SiHO	****	****	****	****	****	****	****	****	****	****
190	SiO	-18.1945	****	****	****	****	****	****	2.2124	-0.0017	0.0278
191	SiH_2	****	****	****	****	****	****	****	-3.6662	****	****
192	SiH	-32.7409	****	****	****	****	****	****	-1.8067	****	****
193	Si	-14.8426	****	****	****	****	****	****	0.8481	-0.0055	0.0196
194	$(CH_3)_3N$	****	****	****	****	****	****	****	****	****	****
195	N=N	****	16.5568	****	****	7.3743	8.2226	8.7799	****	****	****
196	C_{cyclic}=N-	****	****	****	****	****	****	****	-2.3927	-0.0141	0.0369
197	C_{cyclic}=CH-	****	8.1918	4.6627	0.3665	****	****	****	****	0.0094	0.0248
198	C_{cyclic}=NH	****	****	****	****	****	****	****	****	****	****
199	N=O	****	15.8469	****	****	11.5213	11.5358	1.6523	7.7737	0.0123	0.0258
200	C_{cyclic}=C	****	****	****	****	-13.7715	-3.7306	-5.0268	2.7846	0.0009	0.0277
201	P=O	****	****	****	****	-7.1878	3.4676	-5.2535	3.8056	-0.0140	0.0096
202	N=N	****	6.2706	****	****	****	****	****	-0.5829	0.0075	0.0387
203	C=NH	****	****	****	****	****	****	****	****	0.0149	0.0535
204	>C=S	****	****	****	****	1.9104	8.9307	2.2395	-4.2769	0.0019	0.0197
205	aC-CON	****	****	****	****	****	****	****	****	****	****
206	aC=O	****	****	****	****	****	****	****	****	****	****
207	aN-	****	****	****	****	****	****	****	****	****	****
208	-Na	****	****	****	****	****	****	****	-9.7168	****	****
209	-K	****	****	****	****	****	****	****	****	****	****
210	HCONH	****	44.3120	****	****	9.8303	16.8004	13.6895	10.9408	0.0380	0.0190
211	CHOCH	****	****	****	****	****	****	****	****	****	****
212	C_2O	****	****	****	****	****	****	****	****	****	****
213	SiH_3	****	****	****	****	****	****	****	-5.5324	****	****
214	SiH_2O	****	****	****	****	****	****	****	****	****	****
215	CH=C=CH	****	****	****	****	****	****	****	-0.9884	0.0091	0.0345
216	CH=C=C	****	****	****	****	****	****	****	****	****	****
217	OP(=S)O	****	****	****	****	****	****	****	****	****	****
218	R	****	****	****	****	****	****	****	****	****	****
219	$CF_{2cyclic}$	****	4.7102	3.2960	1.2955	2.0667	****	****	-1.5591	0.0087	0.0294
220	CF_{cyclic}	****	-1.0004	-3.0134	-3.8550	-5.8384	****	****	0.6377	-0.0051	0.0121

[a] The symbols F_{p1i}, $\Delta_{vap}H^{o}_{1i}$, $\Delta_{vap}H_{1i}$, $\Delta_{vap}S_{1i}$, δ_{D1i}, δ_{P1i}, δ_{H1i}, ω_{1i}, and V_{m1i} represent the contributions (C_i) of the first-order groups for the corresponding properties.

Table A3. MG method based property models analysed using step-wise regression method: Second-order groups and their contributions [a] for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, and $LogK_{ow}$

	Group	T_{b2j}	T_{c2j}	P_{c2j}	V_{c2j}	T_{m2j}	$\Delta_f G^o_{2j}$	$\Delta_f H^o_{gas\ 2j}$	$\Delta_{fus}H_{2j}$	$LogK_{ow2j}$
1	$(CH_3)_2CH$	0.0071	-0.0371	-0.0011	1.6549	0.0426	-0.3787	-1.1487	0.4162	0.1169
2	$(CH_3)_3C$	0.0121	-0.3488	-0.0018	2.6288	-0.1436	-2.5834	-2.0431	0.2015	0.0573
3	$CH(CH_3)CH(CH_3)$	0.1667	1.8262	-0.0014	-4.1443	0.1640	7.0411	7.6476	-0.5399	0.0193
4	$CH(CH_3)C(CH_3)_2$	0.1860	2.5891	-0.0043	-12.4216	0.0360	10.2099	10.0303	-0.4008	-0.1396
5	$C(CH_3)_2C(CH_3)_2$	0.1902	4.6557	-0.0078	-15.3623	0.9639	14.9779	13.6687	1.5179	****
6	CH_n=CH_m-CH_p=CH_k (k,m,n,p in 0..2)	0.0993	0.5149	-0.0012	-7.5646	0.1319	-6.4113	-9.7132	1.5712	0.0589
7	CH_3-CH_m=CH_n (m,n in 0..2)	0.0111	0.0894	-0.0011	-0.7379	0.0912	0.0097	0.1673	0.3153	0.0699
8	CH_2-CH_m=CH_n (m,n in 0..2)	-0.0638	-0.2233	0.0001	0.8980	-0.1366	1.1234	1.4571	-0.9211	-0.0077
9	CH_p-CH_m=CH_n (m,n in 0..2; p in 0..1)	0.0010	-0.3343	0.0005	-1.6360	-0.1366	3.6828	4.3852	-0.7948	-0.3100
10	CHCHO or CCHO	-0.1363	-2.1746	0.0057	16.5756	0.5854	-1.6156	-0.0237	-3.2688	0.0226
11	CH_3COCH_2	0.0273	-0.0379	0.0010	-2.1274	-0.0205	-0.2548	0.1734	-0.0582	-0.3209
12	CH_3COCH or CH_3COC	-0.1662	-0.6461	0.0001	1.0096	-0.3207	3.3807	5.5072	2.1448	-0.1530
13	CHCOOH or CCOOH	-0.0941	-8.7373	0.0014	10.9060	-0.8078	11.2996	-0.4848	-6.1196	0.2210
14	CH_3COOCH or CH_3COOC	-0.1159	-1.5900	0.0056	4.5375	0.1575	-12.3700	-11.7755	1.6166	0.1958
15	CO-O-CO	0.0869	-1.4141	0.0013	-3.3717	-1.3814	-16.1261	-13.8299	-2.3176	-0.3138
16	CHOH	-0.1193	-1.5693	0.0014	-0.8847	-0.0049	-6.0710	-5.1266	-0.1504	-0.0449
17	COH	-0.2772	-4.7485	0.0052	12.5837	0.4864	-21.4072	-23.7946	-1.4402	-0.3586
18	CH_3COCH_nOH (n in 0..2)	0.0251	-1.2286	-0.0040	3.6856	-0.0909	8.1195	7.3923	****	****
19	NCCHOH or NCCOH	0.3953	0.9065	-0.0021	1.0200	-0.6797	4.8141	11.4510	-0.1134	-0.6779
20	OH-CH_n-COO (n in 0..2)	-0.1171	****	****	****	-0.9076	****	****	****	-0.1516
21	$CH_m(OH)CH_n(OH)$ (m,n in 0..2)	0.1944	1.9395	-0.0014	9.3264	-0.2099	-0.8835	-1.7496	0.5414	-0.1041
22	$CH_m(OH)CH_n(NH_p)$ (m,n,p in 0..2)	0.3136	4.4706	-0.0008	6.7464	-0.1908	-4.9553	-5.1348	0.6844	-0.0975
23	$CH_m(NH_2)CH_n(NH_2)$ (m,n in 0..2)	0.2516	0.3128	0.0001	12.1507	0.3678	0.2413	-3.2743	0.3708	0.4023
24	$CH_m(NH)CH_n(NH_2)$ (m,n in 1..2)	0.1662	4.2200	0.0012	****	****	3.6468	3.9960	-5.7405	0.1997
25	$H_2NCOCH_nCH_mCONH_2$ (m,n in 1..2)	-0.0558	****	****	****	****	****	****	****	-0.1419
26	$CH_m(NH_n)$-COOH (m,n in 0..2)	-0.0614	****	-0.0045	12.2472	1.3964	3.1153	9.6367	8.2846	-0.6678
27	HOOC-CH_n-COOH (n in 1..2)	-0.0771	19.1961	-0.0021	9.7485	0.2553	-5.2337	-2.2874	4.8008	0.3463
28	HOOC-CH_n-CH_m-COOH (n, m in 1..2)	-0.0345	16.2093	0.0018	12.1533	1.1390	-4.3566	-3.1470	29.7631	0.1769
29	HO-CH_n-COOH (n in 1..2)	-0.1538	****	****	****	-0.6687	****	****	****	0.1288
30	NH_2-CH_n-CH_m-COOH (n, m in 1..2)	-0.1656	****	****	****	-9.5780	****	****	****	-1.4198
31	CH_3-O-CH_n-COOH (n in 1..2)	-0.2100	5.1074	-0.0010	7.3876	-1.9485	-1.4197	****	-4.8210	-0.0132
32	HS-CH-COOH	-0.3486	****	****	****	-1.0728	****	****	****	0.1179
33	HS-CH_n-CH_m-COOH (n, m in 1..2)	-0.1109	5.0190	0.0011	16.5075	1.6380	-8.0397	9.0814	-2.5929	0.2316
34	NC-CH_n-CH_m-CN (n, m in 1..2)	-0.0687	17.1845	0.0240	-0.1099	****	13.4079	12.9590	-6.3706	-0.5212
35	OH-CH_n-CH_m-CN (n, m in 1..2)	0.0308	4.9483	0.0095	2.6655	-1.4063	-1.9137	-1.0489	-3.2594	-0.3661

36	HS-CH_n-CH_m-SH (*n, m* in 1..2)	0.1809	6.0898	0.0005	-15.1383	0.6372	5.4773	0.5898	0.6112	****
37	COO-CH_n-CH_m-OOC (*n, m* in 1..2)	0.0799	3.5070	-0.0018	-2.9004	-0.1133	0.0455	4.1113	****	0.0555
38	OOC-CH_m-CH_m-COO (*n, m* in 1..2)	0.2220	4.8206	0.0009	****	0.6655	****	****	2.4237	0.0551
39	NC-CH_n-COO (*n* in 1..2)	0.2657	2.5120	0.0135	0.2995	0.5922	7.8106	****	0.5514	-0.3036
40	$COCH_n$COO (*n* in 1..2)	0.0859	1.8790	-0.0004	1.4473	-0.2616	11.6800	-6.7421	****	0.3100
41	CH_m-O-CH_n=CH_p (*m,n,p* in 0..3)	0.1086	0.6255	-0.0034	-3.0185	0.2622	-11.9809	-10.9343	2.7656	-0.0618
42	CH_m=CH_n-F (*m,n* in 0..2)	0.0763	1.1930	-0.0035	0.0000	0.0034	0.0000	0.0000	0.0000	-0.0295
43	CH_m=CH_n-Br (*m,n* in 0..2)	-0.1303	-3.4098	-0.0121	3.6617	-0.5066	13.4032	16.2949	2.4112	0.0282
44	CH_m=CH_n-I (*m,n* in 0..2)	-0.2420	****	****	****	0.1229	****	****	****	****
45	CH_m=CH_n-Cl (*m,n* in 0..2)	-0.0195	-0.1427	-0.0006	4.4575	-0.0958	5.3462	3.7887	-0.9945	0.1349
46	CH_m=CH_n-CN (*m,n* in 0..2)	-0.0348	-0.5768	0.0059	12.9827	-0.0008	-8.1410	-2.0890	-0.2102	-0.1648
47	CH_n=CH_m-COO-CH_p (*m,n,p* in 0..3)	0.0501	-0.7358	-0.0015	0.6020	0.0776	-11.8853	-11.0525	-0.2889	-0.1289
48	CH_m=CH_n-CHO (*m,n* in 0..2)	0.1708	1.7653	0.0098	****	-0.6543	****	****	****	0.0380
49	CH_m=CH_n-COOH (*m,n* in 0..2)	0.1199	-1.1866	-0.0013	3.8860	0.9794	8.0573	6.6031	2.1022	0.3841
50	aC-CH_n-X (*n* in 1..2) X: Halogen	0.1378	7.6911	0.0033	-10.8937	0.2676	-8.9209	-6.2532	0.3266	0.0747
51	aC-CH_n-NH_m (*n* in 1..2; *m* in 0..2))	-0.0470	2.8668	0.0084	0.6679	-0.3472	-0.2570	-6.8640	-1.3514	-0.3553
52	aC-CH_n-O- (*n* in 1..2)	0.0196	0.0799	0.0012	-2.8197	0.2880	0.6651	1.0429	0.2113	-0.0801
53	aC-CH_n-OH (*n* in 1..2)	0.0894	-0.6470	0.0009	-14.0044	0.5169	-4.5326	-0.7208	0.1511	-0.1083
54	aC-CH_n-CN (*n* in 1..2)	0.0714	10.8883	0.0061	-9.8240	0.0138	24.3938	13.9976	****	-0.6489
55	aC-CH_n-CHO (*n* in 1..2)	0.1664	****	****	****	2.3024	****	****	****	0.4358
56	aC-CH_n-SH (*n* in 1..2)	0.0810	6.4747	0.0095	-9.0889	-0.0485	18.5080	8.5430	0.3371	****
57	aC-CH_n-COOH (*n* in 1..2)	0.0633	6.4527	-0.0008	****	1.1024	****	****	1.6053	0.2284
58	aC-CH_n-CO- (*n* in 1..2)	0.1047	****	****	****	0.1424	****	****	****	-0.2702
59	aC-CH_n-S- (*n* in 1..2)	-0.2178	****	****	****	0.3173	****	****	****	-0.3754
60	aC-CH_n-OOC-H (*n* in 1..2)	0.1730	4.3834	0.0058	-9.8079	1.7051	1.8329	4.9820	****	****
61	aC-CH_m-NO_2 (*n* in 1..2)	0.1486	****	****	****	0.0834	****	-7.5407	****	-0.3072
62	aC-CH_n-$CONH_2$ (*n* in 1..2)	-0.1632	****	****	****	1.5522	****	****	****	-0.2927
63	aC-CH_n-OOC (*n* in 1..2)	-0.1030	3.2086	-0.0014	-14.5368	-0.2286	-0.1204	5.1273	****	0.2754
64	aC-CH_n-COO (*n* in 1..2)	-0.0116	****	****	****	0.3284	****	****	****	-0.1766
65	aC-SO_2-OH	-0.1235	****	****	0.0000	0.5472	0.0000	0.0000	****	-1.3958
66	aC-$CH(CH_3)_2$	0.0073	0.3642	-0.0035	2.5147	-0.0217	2.6148	0.5133	-0.1722	0.0862
67	aC-$C(CH_3)_3$	-0.0199	8.3074	-0.0067	-0.4438	-0.1952	-2.3380	-1.2875	-0.2844	0.1369
68	aC-CF_3	-0.1097	-4.5292	****	****	0.0814	****	****	****	-0.0611
69	(CH_n=C)$_{cyc}$-CHO (*n* in 0..2)	0.1349	6.8126	-0.0011	1.9141	-0.6286	-3.2369	-7.9106	-1.6076	0.4564
70	(CH_n=C)$_{cyc}$-COO-CH_m (*n,m* in 0..3)	0.0897	****	****	****	-0.2405	****	****	****	-0.0048
71	(CH_n=C)$_{cyc}$-CO- (*n* in 0..2)	0.3953	****	****	****	0.9135	****	0.0000	****	0.3321
72	(CH_n=C)$_{cyc}$-CH_3 (*n* in 0..2)	0.0082	-0.5343	-0.0011	0.5568	-0.0482	7.0944	1.0697	0.8756	-0.0654
73	(CH_n=C)$_{cyc}$-CH_2 (*n* in 0..2)	-0.0694	-3.1684	0.0039	-8.2759	-0.7047	-18.0501	-5.9958	-0.1434	-0.2352
74	(CH_n=C)$_{cyc}$-CN (*n* in 0..2)	-0.2645	****	****	****	3.0722	****	****	****	0.0959

75	$(CH_n{=}C)_{cye}$-Cl (*n* in 0..2)	-0.0471	****	****	****	-0.3909	****	****	****	0.1615
76	CH_{cye}-CH_3	-0.0931	-0.3458	-0.0011	-0.7218	-0.1685	3.4844	6.0853	0.2627	-0.0938
77	CH_{cye}-CH_2	-0.0398	0.3439	0.0005	0.5260	-0.2557	-2.0052	2.3177	0.0151	0.1209
78	CH_{cye}-CH	0.0862	0.6262	0.0029	-2.5971	0.2505	-0.2035	7.7340	3.4423	0.3146
79	CH_{cye}-C	-0.0052	3.7347	****	****	0.6051	****	****	****	0.2390
80	CH_{cye}-$CH{=}CH_n$ (*n* in 1..2)	-0.0514	-0.2938	0.0006	-7.0563	-0.3014	2.6589	5.2331	****	0.5075
81	CH_{cye}-$C{=}CH_n$ (*n* in 1..2)	0.0801	****	****	****	-0.9742	****	****	****	0.3580
82	CH_{cye}-Cl	0.0163	-3.0141	****	****	0.6556	****	****	****	-0.0586
83	CH_{cye}-F	-0.0325	****	****	****	3.7390	****	****	****	-0.2523
84	CH_{cye}-OH	-0.1427	2.7089	0.0033	****	0.0821	-7.3373	-9.7727	-6.2753	0.1092
85	CH_{cye}-NH_2	-0.1875	-2.8873	0.0101	****	-0.0910	0.2465	-39.3638	0.6757	-0.2539
86	CH_{cye}-NH-CH_n (*n* in 0..3)	-0.2259	-2.8390	0.0019	****	-0.5454	-17.8392	-14.1356	2.1546	-0.0344
87	CH_{cye}-N-CH_n (*n* in 0..3)	0.0445	****	****	****	0.9660	****	****	****	0.1793
88	CH_{cye}-SH	-0.1105	2.5523	0.0108	****	****	-18.5185	-8.5430	-0.3371	****
89	CH_{cye}-CN	0.4270	****	****	****	0.8050	****	-37.2663	****	-0.0571
90	CH_{cye}-COOH	0.1858	****	****	****	-0.6609	****	****	****	0.3215
91	CH_{cye}-CO	0.1354	****	****	****	0.2376	****	****	****	-0.0480
92	CH_{cye}-NO_2	-0.2571	****	****	****	-0.7987	****	-19.1639	****	-0.0360
93	CH_{cye}-S-	-0.1065	****	****	****	-0.8438	****	****	****	-0.2687
94	CH_{cye}-CHO	-0.0650	****	****	****	0.2252	****	****	****	****
95	CH_{cye}-O-	0.0115	****	****	****	-0.1989	****	****	****	-0.1041
96	CH_{cye}-OOCH	-0.2146	****	****	****	****	****	****	****	0.1912
97	CH_{cye}-COO	0.1121	****	****	****	0.0045	****	****	****	-0.4940
98	CH_{cye}-OOC	-0.2486	-0.4978	-0.0002	****	0.1565	-0.8969	-16.6387	****	0.0580
99	C_{cye}-CH_3	-0.0012	0.1796	-0.0009	1.7750	-0.1837	1.1422	1.3166	-1.3973	0.1140
100	C_{cye}-CH_2	0.0809	0.5885	0.0007	-8.4094	-0.6608	12.3297	9.5703	****	0.1431
101	C_{cye}-OH	-0.2919	-4.0786	0.0015	-13.4138	0.8773	-21.4515	-30.6358	0.5222	0.1790
102	$>N_{cye}$-CH_3	0.0451	-0.0001	-0.0005	0.0000	0.3041	0.0000	-3.2640	****	-0.0925
103	$>N_{cye}$-CH_2	-0.0910	****	****	****	-0.5420	****	-4.0867	****	-0.1025
104	AROMRINGs^{1}s^2	-0.0423	-0.5587	0.0003	6.4951	-0.4527	0.4698	0.4066	-1.7878	-0.0240
105	AROMRINGs^{1}s^3	0.0179	-1.2269	0.0007	10.1641	-0.4331	-1.9640	0.3234	-1.8843	0.1552
106	AROMRINGs^{1}s^4	0.0479	-0.5134	0.0003	6.8997	0.2476	-0.1136	0.0214	0.6993	0.0602
107	AROMRINGs^{1}s^{2}s^3	-0.0054	1.0648	-0.0047	-8.4633	0.1592	3.6947	1.9996	-0.0235	-0.2081
108	AROMRINGs^{1}s^{2}s^4	-0.0003	-0.4645	-0.0031	-10.8758	-0.2140	-1.6151	-2.7023	0.0304	0.0304
109	AROMRINGs^{1}s^{3}s^5	-0.0467	-1.4995	0.0072	-13.1028	0.6065	-2.3129	-9.9138	-0.5689	0.2695
110	AROMRINGs^{1}s^{2}s^{3}s^4	0.0786	5.9314	-0.0067	-18.8966	-0.0618	14.3673	14.7596	-0.3226	-0.1418
111	AROMRINGs^{1}s^{2}s^{3}s^5	-0.0032	2.5689	-0.0025	-11.8966	-0.0113	5.0673	9.5146	-0.8326	-0.1030
112	AROMRINGs^{1}s^{2}s^{4}s^5	0.0453	1.6886	-0.0013	-3.8078	0.7336	-0.3327	6.1346	9.4474	0.1804
113	PYRIDINEs^2	-0.0186	-2.1157	0.0066	7.3722	-0.3354	-10.5206	-2.8843	-10.229	-0.3155

114	PYRIDINEs^3	0.0523	2.1936	0.0066	7.3722	-0.4438	-3.3306	3.0862	-5.7697	-0.5376
115	PYRIDINEs^4	0.1369	2.4148	0.0020	12.9722	0.3046	-5.5106	0.3230	28.4388	-0.4075
116	PYRIDINEs^{2}s^3	0.0557	0.0573	****	-6.6887	-0.0449	****	-0.4596	****	-0.0523
117	PYRIDINEs^{2}s^4	0.0990	-1.6290	0.0042	-6.6887	4.0430	****	-4.3460	****	0.0505
118	PYRIDINEs^{2}s^5	-0.0222	-2.1484	0.0062	-6.6887	0.5361	****	2.0795	****	-0.2864
119	PYRIDINEs^{2}s^6	0.0478	-5.8586	0.0077	****	0.1319	-16.6665	-7.6712	-10.860	-0.0029
120	PYRIDINEs^{3}s^4	0.4669	6.2514	-0.0006	-6.6887	0.3573	****	0.7111	****	-0.3735
121	PYRIDINEs^{3}s^5	0.0900	2.6197	****	****	0.3729	****	2.9556	****	-0.0249
122	PYRIDINEs^{2}s^{3}s^6	-0.0605	****	****	****	****	****	****	****	****
123	(CH^n=CH^m)^{cyc}-COOH	0.1855	****	****	****	2.5389	****	****	****	0.2914
124	AROMRINGs^{1}s^{2}s^{3}s^{4}s^5	-0.0171	****	****	****	-0.1993	****	****	-1.8331	-0.1363
125	aC-NHCOCH$_2$N	-0.1402	****	****	****	-5.1628	****	****	****	-0.3216
126	(N=C)$_{cyc}$-CH$_3$	0.1769	****	****	****	2.4075	****	10.6568	****	-0.2422
127	aC-CONH(CH$_2$)$_2$N	-0.0813	****	****	****	-2.3250	****	****	****	-0.2344
128	aC-SO$_2$NH$_n$ (n>=0;n<3)	-0.0657	****	****	****	0.5634	****	****	****	0.1532
129	aC-SO$_2$NH$_n$ (n>=0;n<3)	-0.0002	****	****	****	-0.3016	****	****	****	0.0250
130	aC-SO$_2$NH$_n$ (n>=0;n<3)	-0.0326	****	****	****	-0.9672	****	****	****	0.1146

[a] The symbols T_{b2j}, T_{c2j}, P_{c2j}, V_{c2j}, T_{m2j}, $\Delta_f G^o_{2j}$, $\Delta_f H^o_{gas2j}$, $\Delta_{fus}H_{2j}$, and $LogK_{ow2j}$ represent the contributions (D_j) of the second-order groups for the corresponding properties

Table A3 (Continued). MG method based property models analysed using step-wise regression method: Second-order groups and their contributions [a] for the properties÷ F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, ω and V_m

	Group	F_{p2j}	$\Delta_{vap}H^o_{2j}$	$\Delta_{vap}H_{2j}$	$\Delta_{vap}S_{2j}$	δ_{D2j}	δ_{P2j}	δ_{H2j}	δ_{2j}	ω_{2j}	V_{m2j}
1	$(CH_3)_2CH$	0.1812	-0.2279	-0.1291	-0.0104	-0.0405	-0.1162	-0.1884	-0.2217	-0.0012	0.0009
2	$(CH_3)_3C$	-3.5999	-0.1046	0.0078	0.0949	0.0266	0.3615	2.3617	-0.0107	-0.0007	0.0012
3	$CH(CH_3)CH(CH_3)$	3.5328	0.8647	1.0810	0.4274	0.2901	-0.5640	-0.0999	0.4609	0.0001	-0.0021
4	$CH(CH_3)C(CH_3)_2$	11.2193	1.4835	1.2189	-0.5534	****	****	****	1.0003	-0.0016	-0.0042
5	$C(CH_3)_2C(CH_3)_2$	19.4129	6.3236	****	****	****	****	****	0.9140	-0.0004	-0.0093
6	CH_n=CH_m-CH_p=CH_k (k,m,n,p in 0..2)	-9.6697	0.8982	-0.2955	-1.1342	0.0869	-0.5470	-1.5359	0.3985	-0.0005	-0.0036
7	CH_3-CH_m=CH_n (m,n in 0..2)	-0.4364	0.6272	0.2071	-0.0106	-0.1801	0.0160	-1.0401	-0.1009	-0.0002	0.0003
8	CH_2-CH_m=CH_n (m,n in 0..2)	-1.6427	-0.0748	0.4031	0.8757	-0.0058	-0.1026	-0.5197	-0.0627	0.0011	-0.0002
9	CH_p-CH_m=CH_n (m,n in 0..2; p in 0..1)	-6.3473	-0.7558	****	****	-0.1039	1.8371	-0.7440	-0.2999	-0.0034	0.0002
10	CHCHO or CCHO	-2.9540	****	****	****	-0.3273	-2.4933	0.8469	-1.2849	0.0031	0.0016
11	CH_3COCH_2	-4.0400	0.1817	-0.3363	-1.1376	-0.0112	-0.7768	-0.3451	-0.4525	-0.0021	0.0002
12	CH_3COCH or CH_3COC	****	1.8337	1.2869	1.8832	-0.2668	-1.5909	-0.5600	-0.0795	0.0008	0.0002
13	CHCOOH or CCOOH	8.4830	5.1503	****	****	0.1079	1.2931	2.4286	0.9532	-0.0031	0.0007
14	CH_3COOCH or CH_3COOC	-24.9284	-1.4900	-1.8607	-1.6599	0.0031	-0.2070	0.1687	-0.2503	-0.0024	0.0030
15	CO-O-CO	5.2878	****	0.2536	0.6467	-0.0963	2.3260	1.4428	0.3818	0.0035	0.0013
16	CHOH	1.0254	1.2650	-1.5532	-0.9922	0.0592	0.3413	0.6216	-0.0472	0.0016	-0.0004
17	COH	-8.6809	0.0799	-1.6892	1.4724	-0.0243	-0.7348	-0.8848	-0.1263	0.0055	0.0015
18	CH_3COCH_nOH (n in 0..2)	****	****	****	****	****	****	****	1.9343	0.0160	-0.0025
19	NCCHOH or NCCOH	0.0000	****	****	****	0.4540	2.8602	-1.1972	1.2624	0.0056	0.0005
20	OH-CH_n-COO (n in 0..2)	-15.3861	****	****	****	0.0072	-0.1667	-3.3827	****	0.0096	-0.0005
21	$CH_m(OH)CH_n(OH)$ (m,n in 0..2)	8.5848	-5.0052	-3.9944	-9.4735	0.1911	-0.8203	-3.2166	-0.1305	-0.0076	0.0015
22	$CH_m(OH)CH_n(NH_p)$ (m,n,p in 0..2)	5.0306	****	1.5721	-2.1116	0.3550	1.2711	-0.8175	1.9880	-0.0016	-0.0006
23	$CH_m(NH_2)CH_n(NH_2)$ (m,n in 0..2)	9.3947	2.6818	1.3580	0.9276	0.2567	-1.7927	3.4033	2.8341	-0.0011	-0.0035
24	$CH_m(NH)CH_n(NH_2)$ (m,n in 1..2)	3.6053	0.4833	-0.2486	-2.1626	0.0605	1.3746	-0.2500	-0.0876	0.0012	0.0011
25	$H_2NCOCH_nCH_mCONH_2$ (m,n in 1..2)	****	****	****	****	****	****	****	****	****	****
26	$CH_m(NH_n)$-COOH (m,n in 0..2)	****	****	****	****	****	****	****	-2.6859	-0.0070	0.0155
27	HOOC-CH_n-COOH (n in 1..2)	****	****	****	****	****	****	****	6.3921	-0.0009	-0.0058
28	HOOC-CH_n-CH_m-COOH (n, m in 1..2)	-2.4277	****	****	****	****	****	****	-2.9731	0.0046	0.0012
29	HO-CH_n-COOH (n in 1..2)	-35.6225	****	****	****	0.6599	0.5382	9.4799	3.7892	0.0183	0.0104
30	NH_2-CH_n-CH_m-COOH (n, m in 1..2)	****	****	****	****	****	****	****	****	****	****
31	CH_3-O-CH_n-COOH (n in 1..2)	****	****	****	****	****	****	****	5.3028	-0.0086	-0.0063
32	HS-CH-COOH	****	****	****	****	-1.0773	1.1718	2.9668	4.3205	-0.0028	-0.0042
33	HS-CH_n-CH_m-COOH (n, m in 1..2)	****	****	****	****	****	****	****	1.9063	-0.0040	-0.0032
34	NC-CH_n-CH_m-CN (n, m in 1..2)	****	****	1.8295	-1.1053	0.5262	-3.6886	1.2713	-0.0650	-0.0013	-0.0010
35	OH-CH_n-CH_m-CN (n, m in 1..2)	****	****	****	****	0.4912	4.0450	2.7001	1.8879	0.0070	-0.0027

36	HS-CH_n-CH_m-SH (*n, m* in 1..2)	****	1.0757	1.6218	1.0697	0.6807	-0.6319	0.2170	1.1975	0.0022	-0.0002
37	COO-CH_n-CH_m-OOC (*n, m* in 1..2)	14.4849	0.1932	****	****	-0.0130	0.5652	0.4053	0.0874	0.0017	0.0013
38	OOC-CH_n-CH_m-COO (*n, m* in 1..2)	****	****	****	****	0.0506	-0.0968	1.1963	0.0129	0.0016	-0.0009
39	NC-CH_n-COO (*n* in 1..2)	****	****	3.8014	6.0436	0.2110	-2.1658	0.6450	1.4017	-0.0045	-0.0013
40	$COCH_nCOO$ (*n* in 1..2)	24.7815	****	****	****	0.1556	-0.6087	0.4474	0.5502	0.0046	-0.0017
41	CH_m-O-CH_n=CH_p (*m,n,p* in 0..3)	-18.0639	1.0948	0.0104	-1.3784	-0.0424	0.0916	0.2445	0.1460	-0.0007	-0.0010
42	CH_m=CH_n-F (*m,n* in 0..2)	****	****	****	****	-0.0886	-0.2863	-4.4333	-0.1972	0.0001	0.0011
43	CH_m=CH_n-Br (*m,n* in 0..2)	****	****	****	****	-0.4889	0.3079	-0.6056	-0.3467	0.0042	-0.0012
44	CH_m=CH_n-I (*m,n* in 0..2)	****	****	****	****	0.1744	-0.8610	0.4795	****	****	****
45	CH_m=CH_n-Cl (*m,n* in 0..2)	2.0414	-0.6037	-0.4149	-0.3691	0.3520	0.5407	-0.7345	-0.0513	-0.0001	-0.0002
46	CH_m=CH_n-CN (*m,n* in 0..2)	-9.0007	0.1546	-0.2467	0.4849	-0.3290	-0.1716	1.6540	0.8807	0.0017	-0.0001
47	CH_n=CH_m-COO-CH_p (*m,n,p* in 0..3)	1.6842	****	0.4156	1.4245	0.2082	0.1368	-0.8720	-0.1609	-0.0020	0.0002
48	CH_m=CH_n-CHO (*m,n* in 0..2)	-5.7993	****	-0.5449	-1.9934	0.5816	0.5027	-0.3584	0.6627	-0.0024	-0.0039
49	CH_m=CH_n-COOH (*m,n* in 0..2)	24.7300	****	****	****	0.5318	0.4454	2.1610	0.9168	-0.0036	-0.0004
50	aC-CH_n-X (*n* in 1..2) X: Halogen	13.4197	****	****	****	0.2213	-0.0846	0.6689	-0.2895	-0.0013	-0.0001
51	aC-CH_n-NH_m (*n* in 1..2; *m* in 0..2))	****	2.0616	****	****	0.4492	-2.2945	1.9559	-2.9980	-0.0030	0.0008
52	aC-CH_n-O- (*n* in 1..2)	0.0000	****	****	****	-0.1365	-0.0158	0.8380	1.5652	0.0012	-0.0002
53	aC-CH_n-OH (*n* in 1..2)	-1.1962	****	****	****	0.5154	0.0164	0.8857	-0.5225	-0.0030	0.0011
54	aC-CH_n-CN (*n* in 1..2)	****	****	****	****	0.6512	0.0623	1.9122	-0.2955	0.0007	0.0014
55	aC-CH_n-CHO (*n* in 1..2)	****	****	****	****	****	****	****	****	-0.0011	-0.0002
56	aC-CH_n-SH (*n* in 1..2)	****	****	****	****	****	****	****	-0.3273	-0.0020	0.0009
57	aC-CH_n-COOH (*n* in 1..2)	****	****	****	****	****	****	****	****	0.0125	0.0188
58	aC-CH_n-CO- (*n* in 1..2)	****	****	****	****	-2.7946	-0.8886	****	****	****	****
59	aC-CH_n-S- (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****
60	aC-CH_n-OOC-H (*n* in 1..2)	****	****	****	****	****	****	****	0.0656	-0.0031	0.0008
61	aC-CH_m-NO_2 (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****
62	aC-CH_n-$CONH_2$ (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****
63	aC-CH_n-OOC (*n* in 1..2)	18.9879	****	-1.1912	-1.6981	-1.0879	-0.3143	0.0979	-0.5926	0.0001	0.0081
64	aC-CH_n-COO (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****
65	aC-SO_2-OH	****	****	****	****	****	****	****	****	****	****
66	aC-$CH(CH_3)_2$	3.1004	0.9546	****	****	0.7291	-1.1559	1.8781	0.7557	0.0004	0.0005
67	aC-$C(CH_3)_3$	0.0000	-2.1016	****	****	-1.4789	-0.0496	1.9725	1.1731	-0.0004	0.0008
68	aC-CF_3	****	2.0992	0.0000	0.0000	0.4225	-0.3328	-8.8009	0.1863	-0.0016	-0.0033
69	$(CH_n{=}C)_{(cyc)}$-CHO (*n* in 0..2)	15.0871	****	2.3557	0.7665	0.5492	2.1202	-3.5531	0.4011	-0.0068	0.0066
70	$(CH_n{=}C)_{cyc}$-COO-CH_m (*n,m* in 0..3)	****	****	****	****	-0.7339	-1.5368	1.5749	****	****	****
71	$(CH_n{=}C)_{cyc}$-CO- (*n* in 0..2)	****	****	****	****	-0.8497	1.9169	2.4173	0.0704	****	****
72	$(CH_n{=}C)_{cyc}$-CH_3 (*n* in 0..2)	-7.4519	-0.2542	-1.8630	-2.6007	-0.0877	0.0667	1.5925	-0.8355	-0.0017	-0.0012
73	$(CH_n{=}C)_{cyc}$-CH_2 (*n* in 0..2)	-7.6915	****	3.2337	7.0348	-1.4776	-2.0653	-0.1157	2.0446	0.0024	0.0063
74	$(CH_n{=}C)_{cyc}$-CN (*n* in 0..2)	****	0.3811	****	****	-0.4273	0.3201	3.0012	****	****	****

75	$(CH_n{=}C)_{cyc}$-Cl (*n* in 0..2)	****	****	****	****	1.6194	-1.1049	2.7380	0.1332	-0.0025	0.0008
76	CH_{cyc}-CH_3	-8.2727	0.1877	-0.2576	-0.6495	-0.7482	-0.0163	-2.6027	0.7560	0.0001	0.0023
77	CH_{cyc}-CH_2	5.6235	-0.1533	-0.1383	-1.3180	-0.0140	-1.7793	-4.0908	-0.2901	0.0022	0.0010
78	CH_{cyc}-CH	12.8835	1.1524	0.9079	0.0952	-1.5719	-1.6180	-1.5266	-0.0797	0.0029	-0.0082
79	CH_{cyc}-C	****	1.9621	0.7445	-0.2722	****	****	****	****	0.0064	0.0056
80	CH_{cyc}-$CH{=}CH_n$ (*n* in 1..2)	5.4702	0.9136	-0.2636	-3.0464	0.6341	0.2930	-3.3474	1.6087	0.0010	0.0041
81	CH_{cyc}-$C{=}CH_n$ (*n* in 1..2)	8.6553	****	****	****	-1.0010	-2.3852	-1.5637	2.8379	0.0146	0.0018
82	CH_{cyc}-Cl	****	****	****	****	-0.9056	1.8420	-1.2717	0.3635	****	****
83	CH_{cyc}-F	****	****	****	****	****	****	****	****	****	****
84	CH_{cyc}-OH	-6.6022	4.3977	3.0890	11.9082	-0.3337	0.8692	-1.3628	-0.9065	0.0023	-0.0033
85	CH_{cyc}-NH_2	-18.1348	-7.0630	-2.6531	-3.2450	-0.9879	-3.1721	-5.1402	-4.1468	0.0024	0.0028
86	CH_{cyc}-NH-CH_n (*n* in 0..3)	****	****	****	****	****	****	****	-0.5744	0.0006	0.0006
87	CH_{cyc}-N-CH_n (*n* in 0..3)	****	****	****	****	****	****	****	****	****	****
88	CH_{cyc}-SH	****	0.0000	0.0000	0.0000	****	****	****	-0.5441	0.0013	0.0022
89	CH_{cyc}-CN	****	-1.1236	0.1070	-1.2269	8.4427	7.0606	3.0634	****	0.0047	0.0061
90	CH_{cyc}-COOH	****	****	****	****	****	****	****	0.1802	0.0008	0.0010
91	CH_{cyc}-CO	****	0.5307	2.0615	1.8182	7.2260	5.7977	11.8459	3.1104	0.0158	-0.0067
92	CH_{cyc}-NO_2	****	****	****	****	****	****	****	****	****	****
93	CH_{cyc}-S-	****	****	****	****	****	****	****	****	-0.0020	0.0031
94	CH_{cyc}-CHO	****	****	****	****	0.4537	-0.2233	-3.5562	0.2006	-0.0013	0.0048
95	CH_{cyc}-O-	****	****	****	****	-0.2910	0.1298	-1.0584	-1.1660	0.0024	0.0053
96	CH_{cyc}-OOCH	****	****	****	****	****	****	****	1.5317	-0.0014	0.0016
97	CH_{cyc}-COO	****	0.0000	-0.6919	-1.7835	-1.0618	-0.0750	-1.1121	****	-0.0042	0.0031
98	CH_{cyc}-OOC	-0.2651	****	****	****	-0.6085	4.0703	-0.0407	2.0339	-0.0060	-0.0357
99	C_{cyc}-CH_3	1.8273	0.0713	-0.0029	-0.0493	0.9088	0.0146	-0.3430	-0.4601	0.0003	0.0029
100	C_{cyc}-CH_2	****	1.0646	-0.7056	-2.3470	****	****	****	-4.0577	0.0025	-0.0025
101	C_{cyc}-OH	-8.1952	****	****	****	-5.6725	-0.8195	-1.6146	-0.4790	-0.0152	-0.0044
102	$>N_{cyc}$-CH_3	0.0000	****	****	****	-1.0144	0.9818	1.9504	1.0236	-0.0017	-0.0048
103	$>N_{cyc}$-CH_2	****	****	****	****	-1.2348	-1.4964	-0.2783	0.5028	0.0030	0.0033
104	AROMRINGs^{1}s^2	-5.3475	1.0186	0.6922	1.3948	0.0310	0.5558	-0.1562	0.0819	-0.0012	-0.0031
105	AROMRINGs^{1}s^3	7.7441	-3.6560	0.1231	0.4984	-0.2746	0.8707	-0.0298	0.5637	0.0002	-0.0010
106	AROMRINGs^{1}s^4	-0.4732	0.4880	0.6051	1.0228	-0.0243	0.2453	-0.4144	-0.2098	0.0005	0.0010
107	AROMRINGs^{1}s^{2}s^3	-1.6525	0.7584	****	****	0.3928	1.4673	-0.5336	-0.3598	0.0014	0.0019
108	AROMRINGs^{1}s^{2}s^4	-2.5265	0.9796	-1.1663	-0.9715	0.0506	0.2714	-0.5098	0.7329	0.0012	0.0005
109	AROMRINGs^{1}s^{3}s^5	17.9811	3.3334	****	****	-0.2583	-0.1508	-0.1894	-0.5972	-0.0006	0.0018
110	AROMRINGs^{1}s^{2}s^{3}s^4	11.6163	****	****	****	0.2637	-0.4967	-0.6207	0.7839	0.0014	-0.0041
111	AROMRINGs^{1}s^{2}s^{3}s^5	6.9493	****	****	****	-0.4168	-1.8048	-2.0195	-2.5308	0.0006	0.0006
112	AROMRINGs^{1}s^{2}s^{4}s^5	****	****	****	****	-0.0507	0.9671	0.4745	-1.5838	0.0022	0.0021
113	PYRIDINEs^2	-3.6285	-2.6133	-1.1926	-1.7315	0.0061	-2.7847	-0.7000	0.5794	0.0045	0.0012

114	PYRIDINEs^{3}	****	-0.5433	-0.0126	-2.0686	****	****	****	-0.2190	0.0012	-0.0001
115	PYRIDINEs^{4}	14.4815	-0.2933	0.1474	-1.9429	-0.0728	1.7325	-0.8632	1.1376	0.0048	0.0001
116	PYRIDINEs^{2}s^{3}	****	-1.3589	-0.0413	-0.7386	****	****	****	****	****	****
117	PYRIDINEs^{2}s^{4}	****	-1.6089	-0.5913	-1.4293	****	****	****	****	****	****
118	PYRIDINEs^{2}s^{5}	****	-1.2689	-0.4413	-0.9986	****	****	****	1.4757	****	****
119	PYRIDINEs^{2}s^{6}	****	-3.7289	-1.6613	-0.9342	****	****	****	0.3481	0.0048	0.0028
120	PYRIDINEs^{3}s^{4}	****	1.4411	0.8687	-2.2870	****	****	****	****	****	****
121	PYRIDINEs^{3}s^{5}	****	0.3911	0.3387	-2.0475	****	****	****	****	****	****
122	PYRIDINEs^{2}s^{3}s^{6}	****	****	****	****	****	****	****	****	****	****
123	(CH^{n}=CH^{m})^{cyc}-COOH	****	****	****	****	****	****	****	****	****	****
124	AROMRINGs^{1}s^{2}s^{3}s^{4}s^{5}	****	0.3776	0.1482	0.1121	****	****	****	****	-0.0007	-0.0070
125	aC-NHCOCH$_2$N	****	****	****	****	****	****	****	****	****	****
126	(N=C)$_{cyc}$-CH$_3$	****	****	****	****	****	****	****	****	****	****
127	aC-CONH(CH$_2$)$_2$N	****	****	****	****	****	****	****	****	****	****
128	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	****	****	****	****	****	****	****	****
129	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	****	****	****	****	****	****	****	****
130	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	****	****	-1.2780	-1.0623	-7.0834	****	****	****

[a] The symbols F_{p2j}, $\Delta_{vap}H^{o}_{2j}$, $\Delta_{vap}H_{2j}$, $\Delta_{vap}S_{2j}$, δ_{D2j}, δ_{P2j}, δ_{H2j}, ω_{2j}, and V_{m2j} represent the contributions (D_j) of the second-order groups for the corresponding properties.

Table A4. MG method based property models analysed using step-wise regression method: Third-order groups and their contributions [a] for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, and $LogK_{ow}$

	Group	T_{b3k}	T_{c3k}	P_{c3k}	V_{c3k}	T_{m3k}	$\Delta_f G^o{}_{3k}$	$\Delta_{fl}H^o{}_{gas3k}$	$\Delta_{fus}H_{3k}$	$LogK_{ow3k}$
1	HOOC-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	13.4712	0.0058	-7.3702	-0.2366	-0.0003	-2.7661	0.0000	-0.1839
2	NH_n-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	-2.3057
3	NH_2-$(CH_n)_m$-OH (m>2, n in 0..2)	****	5.5882	0.0017	-27.799	0.3475	0.0003	3.5589	0.0000	0.0054
4	OH-$(CH_n)_m$-OH (m>2, n in 0..2)	****	4.1384	-0.0006	-6.6795	-0.3779	0.0012	8.0978	0.0000	-0.9089
5	OH-$(CH_p)_k$-O-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****
6	OH-$(CH_p)_k$-S-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****
7	OH-$(CH_p)_k$-NH_x-$(CH_n)_m$-OH (m,k>0; p,n,x in 0..2)	****	****	****	****	****	****	****	****	****
8	CH_p-O-$(CH_n)_m$-OH (m>2; n,p in 0..2)	****	****	****	****	****	****	****	****	****
9	NH_2-$(CH_n)_m$-NH_2 (m>2; n in 0..2))	0.0828	-1.8632	0.0135	-3.2307	0.2640	0.0003	0.7554	0.0000	0.7245
10	NH_k-$(CH_n)_m$-NH_2 (m>2; k in 0..1; n in 0..2)	****	****	****	****	****	****	****	****	****
11	SH-$(CH_n)_m$-SH (m>2; n in 0..2)	0.2655	****	****	****	-0.6280	****	****	****	****
12	NC-$(CH_n)_m$-CN (m>2)	****	16.6171	0.0258	-6.0023	-0.1286	-0.0011	-5.7303	0.0000	-0.7193
13	COO-$(CH_n)_m$-OOC (m>2; n in 0..2)	****	****	****	****	****	****	****	****	****
14	aC-$(CH_n{=}CH_m)_{cyc}$ (fused rings) (n,m in 0..1)	-0.0591	-9.2419	0.0035	-4.1679	0.0719	0.0003	2.0734	0.0000	-0.0329
15	aC-aC (different rings)	0.0273	-2.3599	-0.0133	29.6558	0.0027	-0.0011	-5.1052	0.0000	0.1424
16	aC-CH_{ncyc} (different rings) (n in 0..1)	-0.2130	-2.7789	0.0040	****	-0.2033	-0.0006	-4.6934	0.0000	0.0042
17	aC-CH_{ncyc} (fused rings) (n in 0..1)	-0.0111	-6.1429	-0.0016	-9.1117	-0.2179	-0.0002	8.0750	0.0000	-0.1070
18	aC-$(CH_n)_m$-aC (different rings) (m>1; n in 0..2)	0.4130	11.8299	0.0010	-5.3635	-0.0860	-0.0011	2.1441	12.0237	0.3816
19	aC-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	-0.3323	****	****	****	-1.1281	****	****	****	-0.0805
20	CH_{cyc}-CH_{cyc} (different rings)	-0.0023	6.2025	0.0035	****	0.7559	****	****	****	0.4252
21	CH_{cyc}-$(CH_n)_m$-CH_{cyc}(different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****
22	$CH_{multiring}$	0.1371	1.9976	-0.0033	0.7655	0.2166	-0.0005	-2.9489	0.0000	-0.0419
23	$C_{multiring}$	0.0331	****	****	****	0.0943	****	****	0.0001	-0.0291
24	aC-CH_m-aC (different rings) (m in 0..2)	-0.0343	0.6827	0.0041	-18.346	-0.1199	0.0002	1.4001	0.0000	-0.1034
25	aC-$(CH_m{=}CH_n)$-aC (different rings) (m,n in 0..2)	0.2534	1.8492	0.0025	6.2226	0.8378	-0.4310	-1.3758	3.5569	0.1833
26	$(CH_m{=}C)_{cyc}$-CH=CH-$(C{=}CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	****
27	$(CH_m{=}C)_{cyc}$-CH_p-$(C{=}CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	0.6657
28	aC-CO-aC (different rings)	0.0214	0.5272	****	-13.931	-0.2827	0.0085	13.2669	0.0000	0.1164
29	aC-CH_m-CO-aC (different rings) (m in 0..2)	0.0214	****	****	****	0.0114	****	****	****	0.2238
30	aC-CO-$(C{=}CH_n)_{cyc}$ (different rings) (n in 0..1)	-0.1736	****	****	****	-0.1165	****	****	****	0.2473
31	aC-CO-CO-aC (different rings)	****	****	****	****	2.4653	****	****	0.0000	0.8518
32	aC-CO_{cyc} (fused rings)	0.1142	****	****	****	0.3217	****	32.3581	0.7983	0.0030
33	aC-CO-$(CH_n)_m$-CO-aC (different rings) (m>0; n in 0..2)	****	****	****	****	-0.9900	****	0.1534	****	****

34	aC-CO-$CH_{n\text{cyc}}$ (different rings) (*n* in 0..1)	0.1959	****	****	****	-3.8388	****	****	****	0.6778
35	aC-CO-NH_n-aC (different rings) (*n* in 0..1)	****	****	****	****	0.4986	****	****	0.0000	0.4732
36	aC-NH_nCONH_m-aC (different rings) (*n,m* in 0..1)	0.0000	****	****	****	2.7765	****	****	0.0000	0.1770
37	aC-CO-N_{cyc} (different rings)	****	****	****	****	0.0270	****	****	****	0.2029
38	aC-S_{cyc} (fused rings)	-0.1030	12.1354	0.0072	-19.926	-0.1041	10.8281	14.3386	0.0000	-0.0153
39	aC-S-aC (different rings)	0.2308	****	****	****	0.5232	****	****	****	0.3515
40	aC-PO_n-aC (different rings) (*n* in 0..4)	****	****	****	****	****	****	****	****	****
41	aC-SO_n-aC (different rings) (*n* in 1..4)	****	****	****	****	1.9271	****	****	0.0000	-0.3536
42	aC-$NH_{n\text{cyc}}$ (fused rings) (*n* in 0..1)	0.1296	3.5500	0.0157	****	0.3896	-0.0140	-8.0616	-0.0001	0.0584
43	aC-NH-aC (different rings)	0.1588	5.7738	0.0116	-8.4909	0.3482	-0.0009	-12.488	0.0000	0.2657
44	aC-$(C=N)_{cyc}$ (different rings)	-0.3045	****	****	****	-2.4511	****	****	****	-0.1221
45	aC-$(N=CH_n)_{cyc}$ (fused rings) (*n* in 0..1)	-0.0363	****	****	****	1.0039	****	-32.340	0.0000	0.0181
46	aC-$(CH_n=N)_{cyc}$ (fused rings) (*n* in 0..1)	****	****	****	****	0.5033	****	40.8091	****	-0.6708
47	aC-O-CH_n-aC (different rings) (*n* in 0..2)	****	****	****	****	0.3651	****	****	****	-0.0661
48	aC-O-aC (different rings)	-0.2417	-3.7437	0.0092	-19.169	-1.1809	2.4094	-8.5210	-4.7576	0.1635
49	aC-CH_n-O-CH_m-aC (different rings) (*n,m* in 0..2)	-0.0392	-0.1595	0.0018	5.6364	-0.4040	0.0000	-2.0857	0.0000	-0.3499
50	aC-O_{cyc} (fused rings)	0.0421	5.1557	****	-19.199	0.0048	****	6.6113	0.0000	0.1147
51	AROM.FUSED[2]	0.0166	-3.9457	-0.0016	8.6042	-0.3344	-0.0096	0.8755	0.0000	0.0285
52	AROM.FUSED[2]s^1	-0.0378	2.1630	-0.0035	-19.985	-0.0784	-0.0003	-10.583	0.0000	0.0512
53	AROM.FUSED[2]s^2	-0.0592	-1.1898	-0.0015	-9.4619	0.1479	-0.0008	-4.5841	0.0000	0.0146
54	AROM.FUSED[2]s^2s^3	0.0760	25.0034	****	****	1.0951	****	****	3.4642	-0.0814
55	AROM.FUSED[2]s^1s^4	-0.1731	****	****	****	0.1550	****	****	0.8765	0.0635
56	AROM.FUSED[2]s^1s^2	0.0339	****	****	****	-0.4324	****	****	****	-0.1149
57	AROM.FUSED[2]s^1s^3	-0.1272	26.5956	****	****	1.8481	****	****	****	0.1795
58	AROM.FUSED[3]	0.0454	-2.9209	0.0015	52.7624	0.0564	3.9946	-1.9029	2.1316	0.0218
59	AROM.FUSED[4a]	0.2206	20.4501	0.0138	-36.657	-0.3009	15.0974	11.5161	9.1486	0.0775
60	AROM.FUSED[4a]s^1	0.3508	****	****	****	0.8565	****	23.7178	****	-0.0727
61	AROM.FUSED[4a]s^1s^4	****	****	****	****	-1.5293	****	26.4171	****	0.7898
62	AROM.FUSED[4p]	0.2868	17.0823	0.0065	-36.756	0.7436	-16.074	-23.907	-3.1469	0.1239
63	AROM.FUSED[4p]s^3s^4	****	****	****	****	0.0335	****	****	****	****
64	PYRIDINE.FUSED[2]	-0.0468	2.4430	-0.0024	-0.5352	-0.3160	9.8205	2.4239	-13.756	-0.2553
65	PYRIDINE.FUSED[2-iso]	-0.1104	15.2077	-0.0062	-9.0683	0.2666	0.0042	8.0040	-12.032	-0.2283
66	PYRIDINE.FUSED[4]	0.3738	****	0.0041	-19.571	-0.0560	0.0016	22.9451	0.0000	-0.2366
67	aC-N-CH_{cyc} (different rings)	****	****	****	****	0.9791	****	****	****	-0.0624
68	N multiring	-0.2925	****	****	****	0.6959	****	32.3518	****	-0.2049
69	N_{cyc}-$(CH_2)_3$-N_{cyc} (different rings)	****	****	****	****	-2.2340	****	****	****	-0.0979
70	aC-$COCH_2CH_2$-aC (different rings)	****	****	****	****	****	****	****	****	-0.4613
71	aC-O-$(CH_2)_2$-N_{cyc} (different rings)	****	****	****	****	****	****	****	****	-0.4410

72	aC-CH(OH)$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	-1.1430
73	Ncyc-$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	0.3798
74	aC-$CONHCH_2$-CH_{cyc} (different rings)	****	****	****	****	-0.9475	****	****	****	-0.4323

[a] The symbols T_{b3k}, T_{c3k}, P_{c3k}, V_{c3k}, T_{m3k}, $\Delta_f G^o_{3k}$, $\Delta_f H^o_{gas3k}$, $\Delta_{fus}H_{3k}$, and $LogK_{ow3k}$ represent the contributions (E_k) of the third-order groups for the corresponding properties

Table A4 (Continued). MG method based property models analysed using step-wise regression method: Third-order groups and their contributions [a] for the properties÷ F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, ω and V_m

	Group	F_{p3k}	$\Delta_{vap}H^o_{3k}$	$\Delta_{vap}H_{3k}$	$\Delta_{vap}S_{3k}$	δ_{D3k}	δ_{P3k}	δ_{H3k}	δ_{3k}	ω_{3k}	V_{m3k}
1	HOOC-$(CH_n)_m$-COOH (m>2, n in 0..2)	-11.380	****	****	****	0.1718	2.5721	0.2768	-1.1485	-0.0007	-0.0100
2	NH_n-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	****	****
3	NH_2-$(CH_n)_m$-OH (m>2, n in 0..2)	****	****	****	****	****	****	****	5.5639	0.0114	-0.0048
4	OH-$(CH_n)_m$-OH (m>2, n in 0..2)	27.4803	1.3166	1.5613	-6.2002	0.3289	2.8092	-2.4542	1.9110	0.0025	-0.0019
5	OH-$(CH_p)_k$-O-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****	****
6	OH-$(CH_p)_k$-S-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****	****
7	OH-$(CH_p)_k$-NH_x-$(CH_n)_m$-OH (m,k>0; p,n,x in 0..2)	****	****	****	****	****	****	****	****	****	****
8	CH_p-O-$(CH_n)_m$-OH (m>2; n,p in 0..2)	****	****	****	****	****	****	****	****	****	****
9	NH_2-$(CH_n)_m$-NH_2 (m>2; n in 0..2))	****	3.1010	1.8926	1.8838	****	****	****	-0.4780	-0.0019	0.0175
10	NH_k-$(CH_n)_m$-NH_2 (m>2; k in 0..1; n in 0..2)	****	****	****	****	****	****	****	****	****	****
11	SH-$(CH_n)_m$-SH (m>2; n in 0..2)	****	1.5576	****	****	****	****	****	****	****	****
12	NC-$(CH_n)_m$-CN (m>2)	-2.3623	****	****	****	****	****	****	-1.8323	0.0006	0.0014
13	COO-$(CH_n)_m$-OOC (m>2; n in 0..2)	****	****	****	****	****	****	****	****	****	****
14	aC-$(CH_n{=}CH_m)_{cyc}$ (fused rings) (n,m in 0..1)	****	****	****	****	0.1715	0.7317	0.9507	0.5569	0.0011	-0.0028
15	aC-aC (different rings)	-2.9197	****	****	****	1.4629	1.4222	-1.0851	0.4809	0.0004	0.0020
16	aC-CH_{ncyc} (different rings) (n in 0..1)	****	0.0000	****	****	0.5718	0.1114	-3.5861	0.5787	-0.0010	-0.0001
17	aC-CH_{ncyc} (fused rings) (n in 0..1)	****	-1.0593	0.8445	2.0938	1.1324	-0.3529	0.9306	0.4518	0.0003	-0.0037
18	aC-$(CH_n)_m$-aC (different rings) (m>1; n in 0..2)	8.1165	****	****	****	****	****	****	-1.8809	-0.0052	-0.0037
19	aC-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****	****
20	CH_{cyc}-CH_{cyc} (different rings)	4.4215	****	****	****	-0.5119	3.1317	-9.7602	****	-0.0057	****
21	CH_{cyc}-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****	****
22	$CH_{multiring}$	4.9360	-0.4984	0.0062	-0.0320	1.9595	-0.1324	-0.4329	-0.0863	-0.0013	0.0007
23	$C_{multiring}$	-7.7416	-0.3783	0.0633	0.2622	-0.3218	1.8483	-0.6190	0.3107	0.0005	-0.0026
24	aC-CH_m-aC (different rings) (m in 0..2)	11.7051	****	****	****	-0.5420	0.7394	-2.3539	-0.3447	-0.0001	-0.0013
25	aC-$(CH_m{=}CH_n)$-aC (different rings) (m,n in 0..2)	****	****	****	****	****	****	****	0.0745	0.0004	-0.0043
26	$(CH_m{=}C)_{cyc}$-CH=CH-$(C{=}CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	****	****
27	$(CH_m{=}C)_{cyc}$-CH_p-$(C{=}CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	****	****
28	aC-CO-aC (different rings)	****	****	****	****	-0.1401	1.8062	3.1782	****	-0.0051	-0.0073
29	aC-CH_m-CO-aC (different rings) (m in 0..2)	****	****	****	****	0.0000	0.0000	-1.7604	-1.1174	****	****
30	aC-CO-$(C{=}CH_n)_{cyc}$ (different rings) (n in 0..1)	****	****	****	****	****	****	****	****	****	****
31	aC-CO-CO-aC (different rings)	****	****	****	****	****	****	****	****	****	****
32	aC-CO_{cyc} (fused rings)	-0.5536	****	****	****	0.4543	-0.3313	-0.2357	-1.7803	-0.0021	0.0043
33	aC-CO-$(CH_n)_m$-CO-aC (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****	****

34	aC-CO-$CH_{n\text{cye}}$ (different rings) (n in 0..1)	****	****	****	****	****	****	****	****	****	****
35	aC-CO-NH_n-aC (different rings) (n in 0..1)	****	****	****	****	****	****	****	****	****	****
36	aC-NH_nCONH_m-aC (different rings) (n,m in 0..1)	****	****	****	****	****	****	****	****	****	****
37	aC-CO-N_{cye} (different rings)	****	****	****	****	****	****	****	****	****	****
38	aC-S_{cyc} (fused rings)	****	****	****	****	0.1519	-8.1865	-1.4420	-0.0667	-0.0007	-0.0008
39	aC-S-aC (different rings)	****	****	****	****	****	****	****	****	****	****
40	aC-PO_n-aC (different rings) (n in 0..4)	****	****	****	****	****	****	****	****	****	****
41	aC-SO_n-aC (different rings) (n in 1..4)	****	****	****	****	0.2500	-1.0606	-2.6053	****	****	****
42	aC-$NH_{n\text{cye}}$ (fused rings) (n in 0..1)	****	****	****	****	0.8535	1.1282	0.7173	0.1496	0.0007	-0.0013
43	aC-NH-aC (different rings)	6.4016	****	****	****	-0.2617	-0.2382	-3.5536	-0.7017	-0.0006	0.0039
44	aC-$(C{=}N)_{cye}$ (different rings)	****	****	****	****	****	****	****	****	****	****
45	aC-$(N{=}CH_n)_{cyc}$ (fused rings) (n in 0..1)	****	****	****	****	0.5637	1.5230	0.2899	****	-0.0006	0.0264
46	aC-$(CH_n{=}N)_{cyc}$ (fused rings) (n in 0..1)	****	****	****	****	1.6139	-0.9441	-0.0743	****	****	****
47	aC-O-CH_n-aC (different rings) (n in 0..2)	****	****	****	****	****	****	****	****	****	****
48	aC-O-aC (different rings)	0.0000	****	****	****	-0.3381	-1.1131	-1.3514	-2.2350	-0.0086	0.0017
49	aC-CH_n-O-CH_m-aC (different rings) (n,m in 0..2)	****	****	****	****	0.0663	0.7437	1.9259	3.7528	0.0078	0.0005
50	aC-O_{cye} (fused rings)	****	****	****	****	0.6075	0.0210	-0.5115	-0.5983	-0.0084	0.0074
51	AROM.FUSED[2]	1.2495	1.1282	-1.6888	-4.1879	-1.1001	-0.2317	2.6294	0.3177	0.0014	0.0048
52	AROM.FUSED[2]s^1	-2.5899	-2.0001	3.3778	8.3754	1.0458	0.5653	0.8405	0.1270	-0.0005	-0.0076
53	AROM.FUSED[2]s^2	1.4245	****	****	****	-1.7641	-0.8747	0.0189	-0.3888	0.0004	0.0052
54	AROM.FUSED[2]s^2s^3	****	****	****	****	****	****	****	****	****	****
55	AROM.FUSED[2]s^1s^4	****	2.6018	****	****	****	****	****	****	****	****
56	AROM.FUSED[2]s^1s^2	****	****	****	****	****	****	****	****	****	****
57	AROM.FUSED[2]s^1s^3	****	2.4118	****	****	****	****	****	****	****	****
58	AROM.FUSED[3]	****	****	****	****	****	****	****	-0.3660	-0.0009	0.0055
59	AROM.FUSED[4a]	****	****	****	****	****	****	****	-3.1093	0.0002	-0.0020
60	AROM.FUSED[4a]s^1	****	****	****	****	****	****	****	****	****	****
61	AROM.FUSED[4a]s^1s^4	****	****	****	****	****	****	****	****	****	****
62	AROM.FUSED[4p]	****	-3.6434	****	****	****	****	****	-0.5671	-0.0022	-0.0186
63	AROM.FUSED[4p]s^3s^4	****	****	****	****	****	****	****	****	****	****
64	PYRIDINE.FUSED[2]	-3.4513	1.4209	****	****	0.7694	-1.6426	-0.7278	1.1486	-0.0012	-0.0058
65	PYRIDINE.FUSED[2-iso]	****	-3.1516	****	****	****	****	****	1.6459	-0.0045	-0.0026
66	PYRIDINE.FUSED[4]	****	****	****	****	2.5583	1.6369	-3.0822	-0.4278	-0.0029	0.0077
67	aC-N-CH_{cye} (different rings)	****	****	****	****	****	****	****	****	****	****
68	N multiring	****	****	****	****	1.3837	0.9827	3.2021	-1.1445	0.0011	-0.0001
69	N_{cye}-$(CH_2)_3$-N_{cye} (different rings)	****	****	****	****	****	****	****	****	****	****
70	aC-$COCH_2CH_2$-aC (different rings)	****	****	****	****	****	****	****	****	****	****
71	aC-O-$(CH_2)_2$-N_{cye} (different rings)	****	****	****	****	****	****	****	****	****	****

72	aC-CH(OH)(CH_2)$_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****
73	Ncyc-(CH_2)$_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****
74	aC-CONHCH_2-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****

[a] The symbols F_{p3k}, $\Delta_{vap}H_{3k}{}^{o}$, $\Delta_{vap}H_{3k}$, $\Delta_{vap}S_{3k}$, $\delta_{D3k,}$ $\delta_{P3k,}$ δ_{H3k}, ω_{3k}, and V_{m3k} represent the contributions (E_k) of the third-order groups for the corresponding properties.

Table A5. MG method based property models analysed using simultaneous regression method: First-order groups and their contributions [a] for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$ F_p, and $\Delta_{vap}H^o$

	Group	T_{bli}	T_{cli}	P_{cli}	V_{cli}	T_{mli}	$\Delta_f G^o{}_{li}$	$\Delta_f H^o{}_{gas\ li}$	$\Delta_{fus}H_{li}$	$LogK_{owli}$	F_{pli}	$\Delta_{vap}H^o{}_{li}$
1	CH_3	0.9218	1.0898	0.0100	63.9854	0.7555	-20.4345	-84.0390	1.1420	0.1152	33.0909	1.6141
2	CH_2	0.5780	3.4604	0.0101	56.8278	0.2966	8.1877	-20.6506	2.6516	0.4594	11.4107	4.8014
3	CH	-0.1189	4.6659	0.0107	37.2813	-0.5960	33.2787	37.6525	-0.9254	0.4300	-17.7416	5.7553
4	C	-0.6495	6.6169	0.0075	25.0561	-0.3679	66.0035	96.1819	-2.0279	0.8143	-36.6949	4.9183
5	CH_2=CH	1.4953	5.2031	0.0181	106.7610	1.0430	76.0611	23.7826	1.8927	0.2304	42.4673	-0.4710
6	CH=CH	1.2001	8.2552	0.0194	97.7190	0.6600	95.8888	82.4195	5.4232	0.4297	10.1964	-2.4067
7	CH_2=C	1.0308	7.3554	0.0176	93.6274	0.3327	91.0854	78.1789	3.1782	0.3122	6.0196	-2.9764
8	CH=C	0.7646	10.0135	0.0170	87.2150	-0.3944	114.7548	138.6480	4.3549	0.6026	-15.4082	-5.6340
9	C=C	0.4080	13.5316	0.0209	101.2116	-0.9826	142.3553	201.5682	2.7754	0.5563	-44.0014	-7.2229
10	CH_2=C=CH	2.1764	11.0144	0.0261	140.9994	1.6865	209.8955	161.8935	6.0995	****	****	11.1663
11	CH_2=C=C	1.7104	12.6575	0.0261	148.4107	1.5154	229.8483	213.1922	7.0169	****	****	****
12	C=C=C	****	****	****	****	****	****	****	****	****	****	****
13	CH≡C	1.5883	6.5964	0.0082	81.1724	2.1745	209.8823	187.1424	-1.4111	-0.2171	60.1706	6.3754
14	C≡C	1.2721	11.6977	0.0101	76.6824	1.3372	216.7880	230.2948	5.6768	0.6099	****	11.3710
15	aCH	0.7332	3.7648	0.0046	45.2515	0.6317	19.9241	-0.6732	1.4441	0.1960	18.7143	4.0975
16	aC fused with aromatic ring	1.2531	17.9650	0.0033	4.5530	2.0942	28.3757	20.2301	1.6799	0.2148	18.8824	5.7969
17	aC fused with non-aromatic	1.1611	18.2562	-0.0068	8.4391	1.6734	29.9678	13.1628	1.5330	0.1295	0.5527	5.5248
18	aC except as above	0.8317	21.0902	-0.0001	23.4172	0.9510	36.6009	49.4849	0.2638	0.2487	29.4370	6.3586
19	aN in aromatic ring	1.0507	9.5985	-0.0057	23.7909	2.0185	84.8886	54.1833	4.5449	-0.3810	49.5376	11.8029
20	aC-CH_3	1.2616	7.9731	0.0153	99.8977	1.1155	15.9981	-33.2700	3.0939	0.5432	43.3846	6.9939
21	aC-CH_2	0.8530	11.8624	0.0162	77.4668	-0.1922	43.9629	31.7820	2.3781	0.6971	24.7144	8.9948
22	aC-CH	0.0274	12.1443	0.0183	68.6439	-0.8845	60.0040	82.2207	2.2627	1.0146	-4.5611	9.0081
23	aC-C	-0.6053	7.1924	0.0215	59.4496	-1.2958	103.5315	142.8566	-0.9647	1.3668	-9.6315	13.7692
24	aC-CH=CH_2	1.8860	14.8280	0.0214	110.6358	1.3216	105.9971	64.5454	5.0190	1.2001	60.9284	****
25	aC-CH=CH	1.8784	22.0515	0.0251	100.4258	1.6683	122.8255	123.2564	5.1278	0.7694	41.6313	****
26	aC-C=CH_2	1.4167	16.6802	0.0236	94.0262	0.7823	129.5146	126.2463	4.7422	0.8764	34.6946	****
27	aC-C≡CH	1.8080	16.9729	0.0146	91.0862	1.7154	254.0264	246.7003	3.5090	0.7982	60.5521	****
28	aC-C≡C	2.0510	****	****	****	5.1689	****	****	****	1.1884	****	****
29	OH	2.2476	10.1672	-0.0071	24.4092	3.2424	-161.8492	-213.8185	3.9494	-1.3365	87.6576	23.9705
30	aC-OH	2.5461	24.0543	-0.0100	40.6785	5.1624	-140.7095	-179.2701	9.1184	-0.1125	101.8597	35.6554
31	COOH	3.9741	30.6307	0.0060	78.8678	6.6986	-362.2582	-432.9967	6.4971	-1.2021	128.2657	15.1359
32	aC-COOH	4.5935	43.9567	0.0112	98.4660	14.0994	-321.7557	-379.2337	20.4443	0.0375	163.9821	****
33	CH_3CO	2.6907	14.1929	0.0189	132.0797	3.2535	-149.7355	-213.0782	6.2525	-0.3980	87.7425	12.8497
34	CH_2CO	1.9665	15.1201	0.0178	122.2872	2.8589	-114.0724	-154.6631	8.3352	-0.2250	63.7628	16.7382
35	CHCO	1.1925	14.7945	0.0163	94.3910	1.5650	-90.5148	-91.1938	5.9492	0.2663	****	16.1213
36	CCO	0.5925	****	****	****	1.9597	****	****	****	0.3889	****	14.5913
37	aC-CO	2.2519	29.9384	0.0136	81.0181	2.4974	-79.8262	-85.9307	4.2919	-0.1345	73.3748	22.8491
38	CHO	2.1021	11.2208	0.0045	58.6962	2.9059	-120.2465	-165.7752	9.7013	-0.9687	72.5327	12.7980
39	aC-CHO	2.6711	25.4355	0.0138	98.8795	3.1633	-89.2637	-119.3350	8.4880	-0.1943	91.5521	****
40	CH_3COO	2.6199	12.9819	0.0234	145.4135	2.3902	-323.4199	-421.5642	6.6799	-0.6609	81.5056	20.1649
41	CH_2COO	2.1182	15.6198	0.0223	134.9034	1.4259	-295.3058	-358.6780	8.4006	-0.2896	46.4475	21.0123

42	CHCOO	1.4466	15.3731	0.0238	123.9395	0.9772	-274.3717	-303.2337	7.9878	-0.3449	35.0582	21.8994
43	CCOO	0.5787	****	****	****	0.4546	****	****	****	0.4214	****	18.1851
44	HCOO	2.2249	11.5492	0.0083	90.8410	2.2831	-297.9270	-365.8808	7.3510	-0.9691	66.8687	15.4891
45	aC-COO	2.1193	20.7693	0.0172	88.6432	1.5817	-276.4139	-282.8609	5.1647	-0.0157	54.9432	23.0191
46	aC-OOCH	****	****	****	****	****	****	****	****	-0.4718	****	****
47	aC-OOC	2.0491	23.6563	0.0191	131.0658	3.6401	****	****	7.4161	-0.1516	****	****
48	COO except as above	1.5802	11.7013	0.0112	79.8771	1.3861	-277.0311	-307.1933	9.6822	-0.4116	57.9918	16.8425
49	CH_3O	1.5840	5.6587	0.0133	82.8275	1.5577	-112.7938	-195.3675	4.4580	-0.3882	53.2444	7.4551
50	CH_2O	0.9750	6.2951	0.0152	74.9171	0.6741	-101.6373	-147.7286	4.4603	-0.1205	32.8187	8.8207
51	CH-O	0.3272	5.4851	0.0263	64.3910	-0.0101	-83.3148	-98.9938	5.8192	0.1109	-19.6811	10.2962
52	C-O	-0.3387	7.4881	0.0199	52.4424	-0.2989	-37.5498	-29.5771	2.9553	1.3491	****	9.3276
53	aC-O	1.1476	13.0925	0.0074	31.4549	1.3231	-70.0222	-65.8355	1.2779	0.0450	2.3310	14.0271
54	CH_2NH_2	2.2640	12.9218	0.0068	116.9153	3.3490	47.0463	-49.1973	12.4026	-1.4619	58.6201	14.0998
55	$CHNH_2$	1.4372	11.2221	0.0125	71.5443	36.2974	65.0300	2.6744	5.0882	-1.4109	18.2867	14.5706
56	CNH_2	0.8863	11.0779	0.0142	80.9394	11.5011	85.8747	52.4495	****	-1.2313	****	12.8669
57	CH_3NH	1.9860	10.6020	0.0155	94.4478	2.7394	84.7404	-16.0110	3.2187	-0.8050	54.9961	13.4073
58	CH_2NH	1.2690	9.7528	0.0196	97.5215	2.0378	93.0289	28.5766	7.7184	-0.8698	23.2187	12.3547
59	CHNH	0.5940	8.7773	0.0109	96.3910	1.3226	99.3852	70.2062	1.2992	-0.3477	1.4689	12.4088
60	CH_3N	0.9990	8.1016	0.0192	90.6647	0.8482	121.9323	55.8439	5.6069	-0.3652	48.5699	13.2920
61	CH_2N	0.3324	7.5215	0.0311	82.2100	-0.4084	141.4133	103.5737	2.2815	-0.4084	-13.9652	8.4003
62	aC-NH_2	2.8636	27.9595	0.0020	68.4499	4.3384	58.2977	3.6661	6.0523	-0.3486	118.2386	21.6806
63	aC-NH	1.8969	27.0488	-0.0026	73.5060	2.0755	110.7461	102.2474	2.4701	0.2971	54.6446	15.7481
64	aC-N	1.0680	18.8249	0.0086	96.0454	1.4958	179.8623	170.1188	3.2332	0.3297	25.8875	18.6549
65	NH_2 except as above	1.9867	9.4904	-0.0109	56.7075	3.8058	23.7639	5.1575	6.3670	-1.0653	79.3176	18.4004
66	CH=N	1.1099	****	****	****	9.2631	****	****	****	0.1306	****	****
67	C=N	0.7839	****	****	****	1.3710	****	****	****	0.7815	****	****
68	CH_2CN	3.5608	26.0349	0.0260	147.4795	2.6918	112.9462	53.6676	6.7706	-0.5195	98.3123	21.9015
69	CHCN	2.7070	21.2132	0.0267	128.5271	1.7476	137.1556	112.3365	8.1993	-0.0542	****	23.6917
70	CCN	1.7225	18.0005	0.0291	117.4941	2.8220	150.5683	127.2755	4.1660	-0.0882	43.3055	20.5769
71	aC-CN	3.0904	29.1952	0.0157	98.0862	5.1611	152.9964	151.8327	4.9390	-0.1215	****	****
72	CN except as above	2.7378	14.6309	0.0045	56.4647	3.3709	110.1491	128.0467	5.0569	0.0015	95.8883	21.1206
73	CH_2NCO	2.8854	****	****	****	****	****	****	****	****	****	****
74	CHNCO	2.2381	****	****	****	****	****	****	****	****	****	****
75	CNCO	****	****	****	****	9.8915	****	****	****	****	****	****
76	aC-NCO	2.3798	17.1127	0.0191	134.8280	2.5987	8.1264	-29.1997	****	****	****	****
77	$CH2NO_2$	3.7227	24.3223	0.0146	154.9994	2.7999	****	-102.5905	10.0398	-0.3330	116.3318	29.5463
78	$CHNO_2$	2.8721	24.1207	0.0140	138.5443	0.8539	20.5327	-46.2192	****	-0.0986	80.9275	27.9017
79	CNO_2	2.1309	****	****	****	6.4251	****	-4.2455	-3.7756	0.1837	****	-39.1161
80	aC-NO_2	3.4712	33.7131	0.0123	121.4476	4.5278	43.7819	-31.2082	8.3930	0.0748	120.3647	24.0833
81	NO_2 except as above	2.8709	24.3559	****	****	3.4847	****	-60.9454	5.4098	-0.4689	125.0369	26.2064
82	ONO	1.7330	****	****	****	****	****	****	****	****	****	****
83	ONO_2	2.7558	****	****	****	2.5511	****	-133.8964	8.4812	-0.0640	****	****
84	$HCON(CH_2)_2$	4.4591	****	****	****	****	****	****	****	****	****	36.6521
85	$HCONHCH_2$	5.9314	****	****	****	****	****	****	****	-1.6324	****	46.3963
86	$CONH_2$	5.2882	64.8568	-0.0024	136.4310	11.8818	-147.2546	-236.6971	15.8998	-1.4315	****	****
87	$CONHCH_3$	4.2196	50.9572	0.0160	188.4310	6.5932	-122.6546	-239.9268	16.4898	-0.7471	****	47.9897

88	$CONHCH_2$	5.0238	****	****	****	4.8098	****	-179.2666	****	-0.8403	****	51.4156
89	$CON(CH_3)_2$	4.7548	36.3175	0.0316	242.4310	3.7173	-76.1546	-224.9268	10.5998	-0.5938	153.0369	38.1963
90	$CONCH_3CH_2$	****	****	****	****	1.7487	****	****	****	-0.7131	****	****
91	$CON(CH_2)_2$	3.3830	****	****	****	2.1882	****	****	****	-0.0733	****	38.8180
92	CONHCO	5.7621	****	****	****	9.2737	****	****	****	-1.8911	****	****
93	CONCO	4.0044	****	****	****	3.3303	****	****	****	-0.8259	****	33.8599
94	$aC\text{-}CONH_2$	6.3135	****	****	****	14.8384	****	-176.9197	17.8290	-1.0998	****	****
95	aC-NH(CO)H	5.6658	57.2696	0.0176	141.0862	6.6017	-53.7736	-135.7997	9.6690	-0.5095	****	****
96	aC-N(CO)H	3.7475	****	****	****	3.5580	****	****	****	-0.4768	****	****
97	aC-CONH	5.4326	****	****	****	8.0247	****	****	16.1644	-0.5282	****	****
98	aC-NHCO	5.3177	****	****	****	9.4041	****	****	4.7040	-0.2305	91.5792	****
99	aC-(N)CO	****	****	****	****	6.6280	****	****	****	-0.8181	****	****
100	NHCONH	****	****	****	****	9.7180	****	****	9.4056	-1.3506	****	****
101	NH_2CONH	****	****	****	****	14.2361	****	****	13.6695	-1.3177	****	****
102	NH_2CON	****	****	****	****	11.6648	****	****	10.5200	-0.7922	****	****
103	NHCON	5.1251	****	****	****	7.9386	****	****	12.2240	-0.3600	****	****
104	NCON	2.5898	****	****	****	3.5801	****	46.6202	10.2211	-0.9818	****	****
105	$aC\text{-}NHCONH_2$	4.4036	****	****	****	18.8093	****	****	17.7190	-0.7682	****	****
106	aC-NHCONH	0.7496	****	****	****	20.9688	****	****	21.1744	-0.0756	****	****
107	NHCO except as above	****	****	****	****	4.0641	****	****	****	-0.9979	****	45.5635
108	CH_2Cl	2.2537	11.9698	0.0136	108.2855	1.9463	-45.1420	-110.0475	4.0297	0.4860	61.4828	12.4608
109	CHCl	1.4829	11.5923	0.0138	100.7734	1.1265	-28.7904	-59.5835	****	0.8969	31.7411	10.8717
110	CCl	0.8215	9.0686	0.0130	87.9394	1.3438	-7.1753	-12.4505	-0.9337	0.9630	****	8.1451
111	$CHCl_2$	2.7459	16.9228	0.0191	156.5512	2.4934	-49.6445	-120.2947	6.2575	0.8117	88.8318	17.2831
112	CCl_2	2.1307	****	****	****	3.2470	****	****	1.3969	0.9687	****	18.9721
113	CCl_3	3.1057	18.9929	0.0267	202.4310	4.2883	-63.8646	-142.2268	2.5398	1.5932	****	17.9505
114	CH_2F	1.5154	6.6904	0.0121	85.4310	1.6657	-199.9546	-264.3268	6.1898	-0.1018	****	8.7998
115	CHF	****	5.2651	-0.0269	****	-0.2090	****	****	****	-0.1733	****	****
116	CF	****	****	****	****	****	****	****	****	****	****	****
117	CHF_2	1.6440	6.2516	0.0201	88.5150	1.9149	-430.9546	-500.7268	6.0598	1.0622	****	9.4149
118	CF_2	0.3991	2.3528	0.0096	94.3249	0.3514	****	****	****	0.7829	****	4.0020
119	CF_3	1.1119	1.1505	0.0388	104.5703	1.1548	-630.5845	-708.6621	2.7118	0.7518	****	1.8678
120	CCl_2F	2.1787	9.2817	0.0213	167.3950	3.3365	-272.1993	-341.0621	2.5541	1.5051	****	7.6163
121	HCClF	1.8171	****	****	****	1.2161	****	****	****	0.4974	****	10.8681
122	$CClF_2$	1.5077	5.3324	0.0337	143.0217	1.8010	-416.6797	-502.1936	1.2142	0.9896	****	10.0948
123	aC-Cl	1.6649	16.8708	0.0106	76.7178	1.9911	-8.7977	-27.4387	5.4034	0.8081	65.1890	10.1842
124	aC-F	0.6830	2.8682	0.0099	52.7142	0.9216	-176.9036	-173.3276	1.8430	0.3064	****	4.3095
125	aC-I	2.9439	34.1338	0.0105	110.0862	2.6445	82.3264	84.3003	3.8090	0.8562	****	****
126	aC-Br	2.2390	17.4668	0.0064	76.8793	2.6398	33.9317	28.1830	6.6370	0.9233	80.6631	13.6133
127	-I except as above	2.6522	17.5551	0.0015	100.6734	2.3998	33.5620	19.9974	5.3497	0.5399	****	14.3642
128	-Br except as above	2.0281	11.7182	-0.0045	74.1998	2.0416	-16.2205	-41.4766	3.1165	0.2971	72.8523	9.8075
129	-F except as above	0.7179	-0.3199	0.0132	23.0339	0.8649	-206.2543	-289.2168	3.8168	-0.0005	****	-3.3490
130	-Cl except as above	1.3561	5.2618	0.0038	56.6942	1.4163	-54.4633	-88.5982	4.5154	0.0896	33.3408	6.0183
131	CHNOH	3.7479	****	****	****	4.0756	****	****	****	-0.2921	****	****
132	CNOH	3.1107	****	****	****	4.0968	****	****	****	-0.6688	****	****
133	aC-CHNOH	****	****	****	****	9.4614	****	****	****	-0.1853	****	****

134	OCH_2CH_2OH	3.8181	19.6026	0.0183	155.8593	2.4297	-256.7970	-385.5345	6.8430	-1.1960	131.2719	31.2064
135	$OCHCH_2OH$	3.1220	****	****	****	10.9421	****	****	****	-0.6443	****	****
136	OCH_2CHOH	2.9869	17.9260	0.0185	146.6521	1.7527	-219.5128	-314.0364	11.9899	-0.7404	****	****
137	-O-OH	2.9996	10.9856	-0.0037	52.9544	4.1968	****	-166.8624	****	****	21.8059	****
138	CH_2SH	2.5879	16.1803	0.0115	106.0166	2.1851	6.9435	-46.5497	8.3721	0.3193	66.4735	15.1089
139	CHSH	1.8393	15.0348	0.0092	104.5443	0.5814	31.1527	11.0465	3.0682	-0.0226	****	15.7706
140	CSH	1.1874	15.2220	0.0097	94.9394	4.7552	58.0047	63.0195	-0.5437	0.6687	****	14.8015
141	aC-SH	2.5228	25.5444	0.0070	94.1050	3.0328	39.7264	30.9003	5.5190	0.3323	****	16.6333
142	-SH except as above	2.0774	13.1905	-0.0079	68.4257	2.6727	-3.1011	-19.2417	5.8759	0.1163	****	11.0794
143	CH_3S	2.5029	14.4849	0.0126	117.9143	1.8635	12.2115	-40.4012	6.6582	0.3952	55.8869	14.2894
144	CH_2S	1.9548	16.7752	0.0138	108.2180	0.8881	34.4285	12.9621	3.8234	0.5640	****	15.7534
145	CHS	1.2626	****	****	****	0.5935	****	****	****	0.9271	****	16.6865
146	CS	0.6391	****	****	****	1.5882	****	****	****	1.7367	****	15.2424
147	aC-S-	1.7834	28.7804	0.0093	****	0.9894	****	****	****	0.6474	****	****
148	SO	4.7622	53.1166	-0.0023	84.4107	5.3103	-49.0917	-66.3478	12.9969	-1.1205	144.7965	****
149	SO2	5.0639	48.1596	0.0032	89.2717	6.4215	-255.0761	-301.6848	17.3369	-0.5708	****	****
150	SO_3 (sulfite)	2.8972	21.9226	0.0072	116.7550	****	****	-426.4866	****	0.1555	****	****
151	SO_3 (Sulfonate)	5.0139	****	****	****	5.9382	****	****	****	-1.2153	****	****
152	SO_4 (Sulfate)	4.1394	50.7355	-0.0207	146.0830	3.7901	-516.8267	-616.8872	****	-0.7612	139.0359	****
153	aC-SO	4.3678	****	****	****	4.9397	****	****	****	-0.5474	****	****
154	aC-SO_2	****	****	****	119.9644	4.3705	-306.4516	-345.7812	3.2803	-0.7439	****	****
155	PH (phosphine)	****	****	****	****	****	****	****	****	****	****	****
156	P (Phospine)	0.7096	****	****	****	0.0729	****	****	****	****	****	****
157	PO_3 (Phospite)	1.7858	****	****	****	****	****	****	****	-1.5939	****	****
158	PHO_3 (Phosponate)	4.2814	****	****	****	-0.9347	****	****	****	-2.1734	****	****
159	PO_3 (Phosponate)	2.8787	****	****	****	0.4993	****	****	****	-0.8413	****	****
160	PHO_4 (Phospate)	****	****	****	****	2.3569	****	****	****	-1.9393	****	****
161	PO_4 (Phospate)	2.5427	49.3085	****	91.1666	1.9244	****	-1014.5970	****	-1.5680	88.4741	****
162	aC-PO_4	****	****	****	****	-2.2355	****	-930.1374	8.6616	-0.6473	7.2252	****
163	aC-P	****	158.1177	****	-186.3333	0.1950	****	147.3626	-1.2584	1.0544	****	****
164	CO_3 (Carbonate)	2.2078	16.9427	0.0145	104.0830	2.7606	-434.3517	-513.6866	11.0669	-0.9028	****	20.3300
165	C_2H_3O	2.7140	13.5385	0.0118	122.4399	1.5194	-17.0936	-167.9292	****	-0.6951	74.9195	15.5756
166	C_2H_2O	1.7850	15.4180	0.0077	192.5740	-2.0258	-0.5394	****	****	****	43.4273	****
167	C_2HO	****	****	****	****	-0.3680	****	****	****	0.1716	****	****
168	CH_2 (cyclic)	0.7147	3.6308	0.0068	48.7395	0.5313	8.7292	-29.2709	0.9905	0.1646	17.9616	3.6900
169	CH (cyclic)	0.4137	4.7661	0.0065	38.8249	-0.1298	0.9855	-72.8957	0.8855	0.4600	7.5875	4.8950
170	C (cyclic)	-0.3809	3.6886	0.0081	39.2170	0.1322	24.3604	116.3572	2.1359	0.6366	-24.2854	3.4363
171	CH=CH (cyclic)	1.3771	6.4667	0.0091	83.8235	1.2254	75.1097	37.0718	1.0124	0.2761	26.9856	7.4405
172	CH=C (cyclic)	0.9247	9.1574	0.0087	74.9435	0.8270	63.4709	41.6830	3.7399	0.6105	33.0981	5.9475
173	C=C (cyclic)	0.6624	****	****	****	1.0683	****	****	****	0.8714	****	****
174	CH_2=C (cyclic)	1.2533	10.0641	0.0120	85.4343	0.8039	****	****	4.1053	-0.1397	28.9457	3.9756
175	NH (cyclic)	1.7013	9.1560	-0.0051	48.0002	4.0385	67.3973	24.7863	6.9466	-0.6511	65.6901	16.1331
176	N (cyclic)	1.1003	7.2820	-0.0042	31.1881	0.5618	86.2105	87.1284	-1.2058	-0.4759	49.4879	****
177	CH=N (cyclic)	5.5397	24.3788	0.0353	****	4.9854	****	42.0845	4.1633	-0.2166	****	12.1611
178	C=N (cyclic)	4.9400	****	****	****	5.7227	****	27.9418	****	0.3032	****	****
179	O (cyclic)	0.8846	6.3360	-0.0021	16.3127	1.3472	-123.3599	-147.1515	3.1984	-0.5295	31.0579	6.5706

180	CO (cyclic)	2.2024	33.8819	-0.0025	54.3917	3.3615	-169.6483	-208.3909	4.9552	-0.3100	83.2582	18.2291
181	S (cyclic)	1.8863	11.8977	-0.0024	39.1695	1.7124	-16.8568	-4.2069	4.5015	0.2996	****	13.1644
182	SO_2 (cyclic)	****	68.0095	-0.0014	91.1026	5.7787	-264.2897	-318.0956	8.7491	-0.9653	****	****
183	>NH	****	****	****	****	1.7669	****	****	****	-0.4627	****	****
184	-O-	1.1379	10.6443	0.0011	24.7209	0.6836	-112.0890	****	0.2581	-0.1556	-8.8090	****
185	-S-	1.1954	****	****	****	2.1318	****	****	****	0.7433	****	****
186	>CO	1.0955	****	****	****	1.6089	****	****	****	-0.1157	****	12.4620
187	PO_2	1.3468	****	****	****	****	****	****	****	****	****	****
188	CH-N	0.0212	****	****	****	-0.1632	****	****	****	-0.3074	****	****
189	SiHO	0.0296	****	****	****	-0.8740	****	****	****	****	****	****
190	SiO	-0.3420	****	****	****	-1.4431	****	****	****	0.4349	-43.1666	****
191	SiH_2	0.9301	****	****	****	0.3230	****	****	****	****	****	****
192	SiH	0.2832	****	****	****	0.1946	****	****	****	1.1489	-6.8441	****
193	Si	-0.4468	7.8493	0.0268	****	0.6939	****	****	****	1.9372	-34.2494	****
194	$(CH_3)_3N$	****	****	****	****	****	****	****	****	****	****	****
195	N=N	****	****	****	****	-1.6010	****	181.9486	****	-0.4715	****	15.7447
196	C_{cyclic}=N-	1.2598	****	****	****	0.7097	****	203.4075	****	0.2718	****	****
197	C_{cyclic}=CH-	0.9120	****	****	****	0.1177	****	****	****	0.2950	****	10.7613
198	C_{cyclic}=NH	****	****	****	****	****	****	****	****	-0.4757	****	****
199	N=O	3.0538	****	****	****	1.6530	****	****	****	-0.3601	****	15.2480
200	C_{cyclic}=C	0.4266	17.0856	0.0171	****	0.1965	****	****	****	0.7928	****	****
201	P=O	****	****	****	****	4.2169	****	****	****	-0.5399	****	****
202	N=N	****	****	****	****	0.2488	****	253.6044	****	0.8576	****	7.3987
203	C=NH	****	****	****	****	3.8635	****	****	****	-0.0351	****	****
204	>C=S	****	****	****	****	2.7979	****	99.3628	****	0.5259	****	****
205	aC-CON	****	****	****	****	8.9197	****	****	****	-1.1599	****	****
206	aC=O	****	****	****	****	4.6063	****	****	0.6430	-0.4841	****	****
207	aN-	****	****	****	****	1.7679	****	226.6172	****	-1.7973	****	****
208	-Na	****	****	****	****	****	****	****	****	****	****	****
209	-K	****	****	****	****	****	****	****	****	****	****	****
210	HCONH	****	****	****	****	9.8261	****	****	****	-0.9709	****	44.1463
211	CHOCH	****	****	****	****	0.8353	****	****	****	****	****	****
212	C_2O	****	****	****	****	****	****	****	****	****	****	****
213	SiH_3	****	****	****	****	****	****	****	****	****	****	****
214	SiH_2O	****	****	****	****	****	****	****	****	****	****	****
215	CH=C=CH	****	****	****	****	****	****	****	****	****	****	****
216	CH=C=C	****	****	****	****	****	****	****	****	****	****	****
217	OP(=S)O	****	****	****	****	****	****	****	****	****	****	****
218	R	****	****	****	****	****	****	****	****	****	****	****
219	$CF_{2cyclic}$	0.5117	****	0.0106	****	0.8727	****	****	****	0.5911	****	4.6026
220	CF_{cyclic}	0.4264	****	****	****	-0.8641	****	****	****	0.2909	****	-0.8287

[a] The symbols T_{b1i}, T_{c1i}, P_{c1i}, V_{c1i}, T_{m1i}, $\Delta_f G^o{}_{1i}$, $\Delta_f H^o{}_{gas1i}$, $\Delta_{fus}H_{1i}$, $LogK_{ow1i}$, F_{p1i}, and $\Delta_{vap}H_{1i}$ represent the contributions (C_i) of the first-order groups for the corresponding properties

Table A5 (Continued). MG method based property models analysed using simultaneous regression method: First-order groups and their contributions [a] for the properties÷ $\Delta_{vap}H$, $\Delta_{vap}S$ δ_D, δ_P, δ_H, δ, AiT, ω and V_m

	Group	$\Delta_{vap}H_{1i}$	$\Delta_{vap}S_{1i}$	δ_{D1i}	δ_{P1i}	δ_{H1i}	δ_{1i}	$AiT1_{1i}$	$AiT2_{1i}$	ω_{1i}	V_{m1i}
1	CH_3	2.3797	-0.4840	7.5983	2.3037	2.2105	-1.8029	-0.3516	-57.8605	0.0017	0.0238
2	CH_2	2.3004	0.6423	-0.0023	-0.1664	-0.2150	-0.1323	0.1009	2.6047	0.0019	0.0166
3	CH	1.6577	0.6894	-7.5390	-3.3851	-2.6826	1.0139	0.5829	79.2115	0.0029	0.0084
4	C	0.6634	-0.1297	-15.6455	-5.1979	-6.4821	1.2449	0.6668	95.9781	0.0008	-0.0015
5	CH_2=CH	3.6704	0.3739	7.7504	3.6752	2.7673	-2.1839	-0.2295	-72.0187	0.0027	0.0333
6	CH=CH	3.6489	0.1313	0.4284	3.0492	0.8631	-0.3708	0.0768	-116.5855	0.0046	0.0244
7	CH_2=C	3.2053	0.5815	0.1956	2.3059	1.0623	-0.8300	0.0794	-44.1695	0.0020	0.0230
8	CH=C	4.7812	1.9897	-7.0086	1.2790	-0.1204	1.6192	0.5499	-66.0852	0.0036	0.0108
9	C=C	5.0939	3.0012	-14.6160	-1.4590	-2.9995	2.7550	0.8132	11.7928	0.0007	-0.0021
10	CH_2=C=CH	6.5451	1.2959	7.7662	2.2910	3.6733	-2.7895	****	****	0.0037	0.0431
11	CH_2=C=C	****	****	-0.0965	-2.1075	0.0791	-1.3229	****	****	0.0020	0.0365
12	C=C=C	****	****	****	****	****	****	****	****	****	****
13	CH≡C	5.8019	3.2387	7.7801	3.0384	3.7484	-1.3175	-0.8833	-1078.9102	0.0034	0.0263
14	C≡C	7.6846	4.8196	-0.0503	-0.5528	1.9582	1.0184	****	****	0.0034	0.0159
15	aCH	2.5866	0.1971	3.1133	0.9190	0.7694	-0.0782	-0.1131	-5.1387	0.0013	0.0120
16	aC fused with aromatic ring	4.7372	2.2507	-3.0510	-1.1299	-2.7569	-0.1111	-34.8062	-54.7826	-0.0014	0.0046
17	aC fused with non-aromatic ring	3.4588	1.5553	-4.4881	-0.3554	-2.6898	-1.2301	23.2758	-2.4889	-0.0004	0.0035
18	aC except as above	****	****	-6.6777	-3.0946	-1.4992	-1.2037	-1.9551	-5.7471	0.0016	0.0071
19	aN in aromatic ring	8.0608	9.8797	3.4095	5.9010	3.9266	-0.8197	-0.0832	-62.5438	0.0008	0.0052
20	aC-CH_3	4.1963	-0.8843	3.0811	-1.1846	0.1713	-0.5759	-0.3481	-214.5329	0.0031	0.0302
21	aC-CH_2	4.5211	1.0856	-5.0406	-4.1990	-3.1805	-0.0100	0.1472	-51.0665	0.0042	0.0229
22	aC-CH	****	****	-13.6223	-5.2318	-7.0890	-0.3146	0.5689	19.7360	0.0020	0.0162
23	aC-C	5.6746	3.4909	-21.2544	-8.0888	-7.8370	1.2302	0.6880	-28.5784	0.0015	0.0095
24	aC-CH=CH_2	10.6661	7.6835	3.0147	-2.4385	1.0566	-1.0365	-0.9888	-2293.4743	0.0044	0.0409
25	aC-CH=CH	****	****	-4.6886	-2.0934	-3.6608	0.1657	9.8720	364.5115	0.0056	0.0287
26	aC-C=CH_2	****	****	****	****	****	-0.1801	0.1098	-43.9833	0.0040	0.0297
27	aC-C≡CH	****	****	3.2337	-1.7949	0.1531	-1.5935	-0.0661	-42.2879	0.0009	0.0347
28	aC-C≡C	****	****	-4.3549	****	-2.9946	-0.6637	****	****	0.0007	0.0413
29	OH	17.3709	21.1170	8.0503	5.2379	11.8005	3.0524	-0.1516	32.3056	0.0180	0.0042
30	aC-OH	15.8230	12.8929	3.3399	1.3244	7.7635	1.9054	-0.3393	-153.1902	0.0119	0.0170
31	COOH	****	****	8.4172	3.1400	7.5917	3.6819	0.0375	120.0741	0.0206	0.0207
32	aC-COOH	****	****	3.0852	0.9871	4.5546	2.1528	-0.3866	-293.2920	0.0218	0.0478
33	CH_3CO	8.9922	2.8550	8.1107	6.3823	3.4394	0.4030	0.1335	119.8078	0.0077	0.0347
34	CH_2CO	9.9376	4.9255	0.5371	1.2706	-0.0788	0.7468	0.1214	94.5070	0.0080	0.0283
35	CHCO	9.1129	4.9603	****	****	****	2.5361	****	****	0.0091	0.0199
36	CCO	2.0596	-7.9027	-15.7596	****	****	0.6865	****	****	0.0012	0.0281
37	aC-CO	****	****	-3.8648	0.0858	-0.8306	1.3457	8.4848	260.6371	0.0089	0.0208
38	CHO	8.7662	4.9348	7.8411	7.8726	5.3761	1.4631	0.4254	-30.4259	0.0079	0.0167
39	aC-CHO	****	****	3.7273	3.7488	2.3964	1.2651	6.0646	-33.3858	0.0102	0.0187
40	CH_3COO	13.5533	8.9566	8.1606	2.0405	5.0132	-0.3355	-0.1427	76.0028	0.0111	0.0423
41	CH_2COO	13.1456	9.5467	0.5148	0.2595	2.7824	1.1235	0.1089	110.1176	0.0109	0.0364

42	CHCOO	10.0655	6.4184	-7.2639	-2.8697	-1.1517	2.1116	0.5128	162.2212	0.0087	0.0284
43	CCOO	8.7002	5.7309	-13.1897	-13.7021	-10.5735	****	****	****	0.0092	0.0224
44	HCOO	10.9176	8.7201	7.9230	4.8158	6.8448	-0.1553	-0.3396	-62.1878	0.0086	0.0256
45	aC-COO	15.3658	16.3010	-5.0497	-0.3862	-0.5041	1.3267	0.2576	76.3038	0.0075	0.0312
46	aC-OOCH	****	****	****	****	****	****	****	****	****	****
47	aC-OOC	****	****	-3.6494	-2.0901	-0.0219	1.6265	****	****	0.0091	0.0249
48	COO except as above	10.4710	7.4628	0.4220	2.2043	2.6377	2.4779	0.3262	213.8562	0.0152	0.0196
49	CH_3O	5.8379	1.7348	7.6317	3.2154	3.3464	-1.3922	-0.1637	-64.8653	0.0055	0.0281
50	CH_2O	4.9874	2.8233	0.1706	0.5137	0.8246	0.0991	0.3876	15.7001	0.0064	0.0228
51	CH-O	4.0605	3.6325	-7.6174	-2.7093	-2.1543	0.5968	0.3381	82.8063	0.0030	0.0205
52	C-O	2.8917	1.8475	****	****	****	2.1448	****	****	0.0037	0.0049
53	aC-O	8.5679	6.9671	-4.4171	-1.2975	0.0149	1.8991	1.3612	-25.1637	0.0068	0.0084
54	CH_2NH_2	9.6773	5.7817	8.1995	5.2101	6.7984	0.3355	0.0522	36.9000	0.0104	0.0262
55	$CHNH_2$	8.1069	7.0637	-0.3812	0.5616	2.8953	1.8389	50.4249	235.1768	0.0101	0.0214
56	CNH_2	8.1526	10.6406	****	****	****	0.2906	****	****	0.0053	0.0144
57	CH_3NH	9.4451	7.0907	7.7307	2.5065	7.2551	-0.2982	34.8042	220.7430	0.0071	0.0279
58	CH_2NH	6.6499	3.3224	0.0223	-0.7159	1.4183	0.2836	0.2006	14.3974	0.0096	0.0246
59	CHNH	5.0623	2.7686	-7.5377	-4.6694	-2.2824	1.1362	0.7001	67.2698	0.0091	0.0182
60	CH_3N	3.2572	1.9902	0.3494	-0.4250	2.4585	-0.4539	13.4754	57.8333	0.0049	0.0265
61	CH_2N	5.2316	5.2118	-6.7232	-0.7354	-7.3014	0.0284	1.2799	120.7779	0.0067	0.0190
62	aC-NH_2	13.1059	6.6437	4.3726	1.9767	6.3616	1.9769	-0.1176	9.8937	0.0096	0.0260
63	aC-NH	****	****	-3.9623	-2.6745	3.2590	0.2889	-0.0471	-212.9195	0.0089	0.0118
64	aC-N	****	****	****	****	****	1.4103	1.0105	132.0401	0.0049	0.0190
65	NH_2 except as above	11.6741	9.8812	9.0936	6.3555	9.0778	5.1519	0.2561	-2.1921	0.0066	0.0096
66	CH=N	****	****	****	****	****	****	****	****	****	****
67	C=N	****	****	-8.1282	-3.2229	-3.1673	-5.1194	****	****	****	****
68	CH_2CN	14.2018	2.7796	8.5643	10.8264	3.3144	3.0873	-0.0555	146.0208	0.0123	0.0314
69	CHCN	13.0032	3.1950	1.3130	4.1991	1.7584	4.4365	****	****	0.0116	0.0254
70	CCN	10.6693	2.2812	-7.7575	-0.8832	0.4757	4.0573	0.2795	2.3359	0.0069	0.0211
71	aC-CN	17.8661	14.1320	3.4370	4.8030	0.6067	1.3967	****	****	0.0085	0.0294
72	CN except as above	12.3216	1.1962	8.1538	9.6317	1.9219	1.2762	-0.3003	9.1856	0.0071	0.0158
73	CH_2NCO	****	****	7.8017	9.6963	0.2895	****	****	****	0.0122	0.0402
74	CHNCO	****	****	****	****	****	****	****	****	****	****
75	CNCO	****	****	****	****	****	****	****	****	****	****
76	aC-NCO	****	****	3.4988	1.1837	1.8845	0.0056	****	****	0.0092	0.0325
77	$CH2NO_2$	19.6283	13.0219	8.9178	11.3660	3.1398	0.8521	59.6801	194.5951	0.0131	0.0330
78	$CHNO_2$	17.4132	10.8352	0.8885	6.5226	0.8294	3.5365	0.1970	42.1198	0.0134	0.0249
79	CNO_2	-24.0533	-4.1948	-6.8570	0.3445	-0.5293	-4.4203	****	****	-0.0083	0.0504
80	aC-NO_2	****	****	3.7302	2.8689	1.2032	1.1883	0.5271	173.0964	0.0123	0.0308
81	NO_2 except as above	16.5635	7.4912	7.5368	3.8857	2.9346	0.8944	-0.0279	13.3422	0.0132	0.0180
82	ONO	****	****	****	****	****	****	****	****	****	****
83	ONO_2	****	****	8.0032	7.4830	2.7871	0.1194	2.1376	-22.3503	0.0157	0.0280
84	$HCON(CH_2)_2$	****	****	0.7580	5.8753	1.6091	4.5168	****	****	****	****
85	$HCONHCH_2$	****	****	9.6017	7.6963	11.7895	9.2857	****	****	****	****
86	$CONH_2$	****	****	9.0888	9.0329	10.6398	11.2741	****	****	0.0148	0.0147
87	$CONHCH_3$	****	****	7.4017	1.9963	3.4895	8.0908	****	****	0.0160	0.0380

88	$CONHCH_2$	****	****	****	****	****	8.0168	****	****	****	****
89	$CON(CH_3)_2$	****	****	9.0040	8.9127	6.8046	3.2384	0.0473	15.4733	0.0124	0.0533
90	$CONCH_3CH_2$	****	****	****	****	****	****	****	****	****	****
91	$CON(CH_2)_2$	****	****	-6.3948	4.3888	0.8686	5.0479	****	****	****	****
92	CONHCO	****	****	****	****	****	****	****	****	****	****
93	CONCO	****	****	****	****	****	****	****	****	****	****
94	$aC\text{-}CONH_2$	****	****	5.6337	10.1051	7.3531	****	****	****	0.0175	0.0349
95	aC-NH(CO)H	****	****	****	****	****	4.3137	****	****	0.0168	0.0353
96	aC-N(CO)H	****	****	****	****	****	****	****	****	****	****
97	aC-CONH	****	****	****	****	****	****	****	****	****	****
98	aC-NHCO	****	****	-4.0045	4.0298	3.8028	1.8432	-0.3394	-924.4939	0.0179	0.0326
99	aC-(N)CO	****	****	****	****	****	3.8786	****	****	****	****
100	NHCONH	****	****	****	****	****	****	****	****	****	****
101	NH_2CONH	****	****	10.9112	5.5671	8.1602	****	****	****	****	****
102	NH_2CON	****	****	****	****	****	****	****	****	****	****
103	NHCON	****	****	****	****	****	****	****	****	****	****
104	NCON	****	****	-13.6930	-1.0149	2.1581	8.3790	****	****	****	****
105	$aC\text{-}NHCONH_2$	****	****	****	****	****	****	****	****	****	****
106	aC-NHCONH	****	****	****	****	****	****	****	****	****	****
107	NHCO except as above	****	****	****	****	****	****	****	****	****	****
108	CH_2Cl	9.1122	3.1362	8.8569	4.3288	2.1489	-0.0226	-0.3251	-7.3040	0.0041	0.0317
109	CHCl	6.1613	1.0093	0.6805	2.4897	-0.5733	-0.0418	-1.1691	-4630.9257	0.0036	0.0267
110	CCl	3.2834	0.4579	-7.3725	0.7705	-6.9931	0.7536	****	****	0.0015	0.0188
111	$CHCl_2$	11.3784	3.5583	9.2349	3.3634	2.5206	-0.4749	23.7818	279.1836	0.0043	0.0448
112	CCl_2	****	****	1.0669	1.0329	-1.4311	0.5181	****	****	****	****
113	CCl_3	12.3064	2.4613	9.9021	2.0681	0.1333	-1.9254	-0.3384	-6.9506	0.0037	0.0621
114	CH_2F	9.1503	10.1686	7.3856	4.3614	-1.2477	-0.7769	****	****	0.0050	0.0251
115	CHF	****	****	-0.0751	-1.0634	-0.2549	****	****	****	-0.0027	0.0193
116	CF	****	****	****	****	****	****	****	****	****	****
117	CHF_2	****	****	7.6277	6.0103	1.0771	-2.7521	****	****	0.0050	0.0349
118	CF_2	2.3658	0.7630	-0.7251	****	****	-0.5348	****	****	0.0007	0.0252
119	CF_3	1.0289	-0.6317	7.8128	3.1133	2.0149	-3.7366	****	****	0.0095	0.0412
120	CCl_2F	4.2921	-2.2831	7.0790	-1.0631	****	-3.1637	****	****	0.0040	0.0567
121	HCClF	6.9857	2.3493	7.7475	0.3389	3.8252	-2.6893	-0.2184	32.3070	0.0035	0.0398
122	$CClF_2$	7.4432	2.6054	6.5642	1.2526	****	-3.6482	-0.4535	-18.0351	0.0044	0.0546
123	aC-Cl	5.9998	-0.2357	3.6334	0.5595	0.6238	-0.1939	-0.3539	-188.8730	0.0030	0.0258
124	aC-F	2.7989	0.9716	2.9653	0.4144	0.2767	-0.6712	****	****	0.0026	0.0167
125	aC-I	****	****	3.9337	1.4051	2.2531	0.0471	****	****	0.0019	0.0361
126	aC-Br	****	****	4.0118	1.7621	1.6744	-0.0395	-0.1804	-59.0137	0.0026	0.0300
127	-I except as above	9.7527	-4.1295	9.2913	2.8794	3.7401	0.5685	****	****	0.0018	0.0246
128	-Br except as above	6.5560	0.9157	9.2791	3.3944	2.9788	-0.4249	-0.4514	-94.6302	0.0020	0.0213
129	-F except as above	-0.3582	-0.4755	6.4464	4.7756	5.1354	-1.2592	-0.6446	-173.0902	0.0017	0.0123
130	-Cl except as above	6.2737	2.5943	8.0525	2.7611	1.8748	-0.7342	-0.5500	-151.2099	0.0018	0.0181
131	CHNOH	****	****	8.7017	1.6963	17.9895	****	****	****	0.0262	0.0227
132	CNOH	****	****	0.3046	-0.2243	5.0366	****	****	****	****	****
133	aC-CHNOH	****	****	****	****	****	****	****	****	****	****

134	OCH_2CH_2OH	21.6722	14.5475	8.3027	5.7113	9.9562	2.3257	47.4817	30.5117	0.0213	0.0417
135	$OCHCH_2OH$	****	****	0.3121	-1.5734	5.1167	****	****	****	0.0279	0.0381
136	OCH_2CHOH	****	****	0.0881	0.7182	4.8512	4.2940	****	****	0.0225	0.0392
137	-O-OH	****	****	7.4017	12.6963	27.7895	5.2932	****	****	0.0257	0.0104
138	CH_2SH	9.9537	2.4892	8.5494	3.8374	4.2415	-0.4082	-0.0764	-87.3884	0.0042	0.0341
139	CHSH	8.4819	2.4729	1.3616	1.9228	2.2675	0.0341	****	****	0.0035	0.0289
140	CSH	6.7795	1.2260	****	****	****	0.3420	****	****	0.0021	0.0216
141	aC-SH	11.8961	5.4663	4.4582	0.0702	4.5741	0.7639	****	****	0.0027	0.0269
142	-SH except as above	8.2684	2.2265	8.7339	2.5118	6.0326	0.8350	****	****	0.0022	0.0161
143	CH_3S	9.6582	2.4728	8.8002	3.6190	3.2506	-0.5579	10.4369	11.5754	0.0037	0.0356
144	CH_2S	9.1184	2.8496	0.7294	-0.0130	0.1251	0.8547	****	****	0.0047	0.0294
145	CHS	7.8976	3.7003	****	****	****	****	****	****	****	0.0717
146	CS	4.7265	-0.3659	****	****	****	****	****	****	0.0036	0.0133
147	aC-S-	****	****	-3.5646	-2.0986	-1.3574	****	****	****	0.0053	0.0317
148	SO	****	****	1.7892	7.4639	3.7447	9.5655	6.3266	78.1180	0.0019	0.0143
149	SO2	****	****	1.6660	8.4335	-0.3325	4.2270	****	****	0.0070	0.0237
150	SO_3 (sulfite)	****	****	****	****	****	2.0409	****	****	0.0141	0.0314
151	SO_3 (Sulfonate)	****	****	0.8540	2.5127	2.8046	****	****	****	-0.0042	0.0552
152	SO_4 (Sulfate)	****	****	1.5058	11.4089	4.1941	4.1431	7.2272	185.8736	0.0074	0.0341
153	aC-SO	****	****	-3.7969	6.4368	-1.0695	****	****	****	****	****
154	aC-SO_2	****	****	-3.2918	9.0285	-0.3191	-11.2009	****	****	****	****
155	PH (phosphine)	****	****	****	****	****	****	****	****	****	****
156	P (Phospine)	****	****	-5.7575	-4.6832	****	****	****	****	0.0024	0.0192
157	PO_3 (Phospite)	****	****	****	****	****	2.5231	****	****	****	****
158	PHO_3 (Phosponate)	****	****	****	****	****	8.4168	****	****	****	****
159	PO_3 (Phosponate)	****	****	-6.0459	5.4504	0.6291	****	****	****	****	****
160	PHO_4 (Phospate)	****	****	****	****	****	****	****	****	****	****
161	PO_4 (Phospate)	****	****	-6.3010	4.3366	2.1176	3.6003	0.1377	74.6000	0.0107	0.0314
162	aC-PO_4	****	****	-13.4419	1.7377	-1.7608	-0.1182	4.8253	66.4905	0.0190	0.0609
163	aC-P	****	****	****	****	****	1.3236	****	****	-0.0039	0.0346
164	CO_3 (Carbonate)	****	****	0.1058	0.6589	2.3941	2.2994	****	****	0.0124	0.0231
165	C_2H_3O	10.5281	6.8989	9.5480	4.6184	7.9801	0.9020	-0.8075	-210.2024	-0.0038	0.0285
166	C_2H_2O	****	****	4.6310	4.0936	2.0651	0.2605	-2.2121	-73.3599	0.0016	0.0290
167	C_2HO	****	****	****	****	****	****	****	****	****	****
168	CH_2 (cyclic)	2.3438	-0.0399	2.6693	0.5104	0.6159	-0.3085	0.0198	-3.1112	0.0009	0.0160
169	CH (cyclic)	2.2834	0.6520	-3.6903	-3.2065	-0.5171	-0.6304	0.0719	15.8622	-0.0078	0.0059
170	C (cyclic)	3.0140	4.0775	-10.2229	-2.6202	-2.6329	1.4521	-0.5077	79.6787	-0.0002	0.1393
171	CH=CH (cyclic)	4.8318	1.2589	7.1860	0.7482	1.9775	-0.7844	-0.4279	-49.4820	0.0024	0.0246
172	CH=C (cyclic)	6.2169	4.4891	-1.1666	3.3074	-1.5025	1.1486	0.2903	33.1340	0.0059	0.0020
173	C=C (cyclic)	****	****	-7.5565	-7.2796	-17.0120	-1.9522	****	****	0.0127	-0.0361
174	CH_2=C (cyclic)	-5.4822	-21.8222	7.2048	4.2872	2.4163	-0.0635	4.6088	-32.6637	0.0040	0.0313
175	NH (cyclic)	10.2414	7.9218	4.9384	2.5434	3.4090	1.8061	-0.3537	-51.9576	0.0047	0.0035
176	N (cyclic)	****	****	-3.1834	-0.6202	-0.6760	0.4942	0.4415	-47.7728	0.0000	0.0017
177	CH=N (cyclic)	8.3067	5.6143	7.5704	10.7679	6.2565	-3.0842	****	****	0.0022	0.0220
178	C=N (cyclic)	****	****	****	****	****	****	****	****	0.0021	****
179	O (cyclic)	4.4546	3.0407	4.1019	3.1910	2.8020	0.9963	-0.3026	-111.5440	0.0040	0.0020

180	CO (cyclic)	20.3720	27.1918	3.7195	6.5431	2.7658	3.1548	12.9173	198.8266	0.0091	0.0120
181	S (cyclic)	9.1865	2.9614	5.4726	3.0043	2.9683	1.6101	****	****	0.0013	0.0101
182	SO_2 (cyclic)	****	****	7.9131	11.9870	4.9204	4.9452	****	****	0.0094	0.0164
183	>NH	****	****	****	****	****	2.3229	****	****	0.0084	0.0061
184	-O-	****	****	-0.7007	-3.1505	0.2654	-3.7715	0.0998	88.3361	-0.0014	0.0104
185	-S-	****	****	0.3194	-0.5306	-0.1864	****	****	****	0.0052	0.0157
186	>CO	****	****	0.1063	1.0367	1.6226	-1.5646	****	****	0.0111	0.0128
187	PO_2	****	****	****	****	****	****	****	****	****	****
188	CH-N	****	****	****	****	****	****	****	****	****	****
189	SiHO	****	****	****	****	****	****	****	****	****	****
190	SiO	****	****	****	****	****	1.8263	0.8070	145.5876	0.0020	0.0283
191	SiH_2	****	****	****	****	****	-3.3919	8.9508	93.3669	****	****
192	SiH	****	****	****	****	****	-1.9389	1.1989	189.5422	****	****
193	Si	****	****	****	****	****	0.3123	0.9090	113.5107	0.0008	0.0221
194	$(CH_3)_3N$	****	****	****	****	****	****	****	****	****	****
195	N=N	****	****	7.7650	8.1102	8.7799	****	****	****	****	****
196	C_{cyclic}=N-	****	****	****	****	****	-2.3272	****	****	-0.0037	0.0369
197	C_{cyclic}=CH-	4.8627	0.6920	****	****	****	****	****	****	0.0053	0.0269
198	C_{cyclic}=NH	****	****	****	****	****	****	****	****	****	****
199	N=O	****	****	11.1114	11.1997	1.6523	8.1503	****	****	0.0033	0.0258
200	C_{cyclic}=C	****	****	-13.7208	-4.6294	-5.0268	3.6338	****	****	0.0036	-0.0006
201	P=O	****	****	-6.8355	2.2530	-5.2535	3.6955	****	****	-0.0051	0.0089
202	N=N	****	****	****	****	****	0.0376	****	****	0.0032	0.0387
203	C=NH	****	****	****	****	****	****	****	****	0.0086	0.0535
204	>C=S	****	****	1.3419	7.7242	2.2395	-5.6389	****	****	0.0014	0.0197
205	aC-CON	****	****	****	****	****	****	****	****	****	****
206	aC=O	****	****	****	****	****	****	****	****	****	****
207	aN-	****	****	****	****	****	****	****	****	****	****
208	-Na	****	****	****	****	****	-9.4455	-0.5718	-91.0759	****	****
209	-K	****	****	****	****	****	****	****	****	****	****
210	HCONH	****	****	9.8017	16.4963	13.6895	11.3879	****	****	0.0141	0.0194
211	CHOCH	****	****	****	****	****	****	****	****	****	****
212	C_2O	****	****	****	****	****	****	****	****	****	****
213	SiH_3	****	****	****	****	****	-5.0540	8.9664	-64.2746	****	****
214	SiH_2O	****	****	****	****	****	****	****	****	****	****
215	CH=C=CH	****	****	****	****	****	-0.7141	****	****	0.0035	0.0352
216	CH=C=C	****	****	****	****	****	****	****	****	****	****
217	OP(=S)O	****	****	****	****	****	****	****	****	****	****
218	R	****	****	****	****	****	****	****	****	****	****
219	$CF_{2cyclic}$	3.3419	1.2815	2.0667	****	****	-1.4072	****	****	0.0029	0.0294
220	CF_{cyclic}	-3.1022	-4.1327	-5.7462	****	****	0.8010	****	****	-0.0016	0.0121

[a] The symbols $\Delta_{vap}H_{li}$, $\Delta_{vap}S_{li}$, δ_{Dli}, δ_{Pli}, δ_{Hli}, Ait_{a1i}, Ait_{b1i}, ω_{1i}, and V_{m1i} represent the contributions (C_i) of the first-order groups for the corresponding properties.

Table A6. MG method based property models analysed using simultaneous regression method: Second-order groups and their contributions [a] for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p and $\Delta_{vap}H^o$

	Group	T_{b2j}	T_{c2j}	P_{c2j}	V_{c2j}	T_{m2j}	$\Delta_f G^o{}_{2j}$	$\Delta_f H^o{}_{gas\ 2j}$	$\Delta_{fus}H_{2j}$	$LogK_{ow2j}$	F_{p2j}	$\Delta_{vap}H^o{}_{2j}$
1	$(CH_3)_2CH$	0.0563	-0.0371	0.0005	6.8664	0.1542	-0.9870	-2.8343	1.4882	0.2753	0.0189	-0.2324
2	$(CH_3)_3C$	0.0460	-0.3488	0.0024	5.4499	-0.1090	-4.6712	-4.1684	0.9932	0.2401	-9.5710	1.4796
3	$CH(CH_3)CH(CH_3)$	0.2237	1.8262	-0.0024	-1.5938	0.3071	7.3275	6.8223	0.0542	0.2673	4.7295	1.0048
4	$CH(CH_3)C(CH_3)_2$	0.2326	2.5891	-0.0055	-12.0589	0.1512	9.6985	8.7902	0.4109	0.0984	9.0931	2.1608
5	$C(CH_3)_2C(CH_3)_2$	0.2048	4.6557	-0.0095	-15.7314	1.0283	14.9084	13.1321	2.5250	****	11.9831	7.4619
6	CH_n=CH_m-CH_p=CH_k (k,m,n,p in 0..2)	0.0403	0.5149	-0.0007	-6.5956	0.4981	-11.1659	-22.6343	0.1644	0.3190	-7.8829	11.5967
7	CH_3-CH_m=CH_n (m,n in 0..2)	-0.0063	0.0894	0.0013	0.0378	0.2949	-0.1821	-4.5070	0.1117	0.2466	1.7685	5.7144
8	CH_2-CH_m=CH_n (m,n in 0..2)	-0.1057	-0.2233	0.0004	1.6785	-0.0881	-0.2330	-3.8094	-1.8289	0.1252	1.6669	5.3637
9	CH_p-CH_m=CH_n (m,n in 0..2; p in 0..1)	-0.0195	-0.3343	0.0000	-0.3479	0.0257	1.8425	-1.8562	-1.2918	-0.1533	-5.6014	5.2250
10	CHCHO or CCHO	-0.0567	-2.1746	0.0054	21.3447	0.7062	-2.9901	-2.7467	-3.7705	0.0393	-2.2503	****
11	CH_3COCH_2	-0.0384	-0.0379	-0.0012	-7.4989	-0.1742	6.3808	-5.6940	0.3783	-0.4799	-11.1798	2.0680
12	CH_3COCH or CH_3COC	-0.2119	-0.6461	-0.0026	-4.9883	-0.4152	9.9892	-0.7559	2.7243	-0.2698	****	3.8571
13	CHCOOH or CCOOH	-0.0330	-8.7373	0.0007	22.7237	-0.2224	15.1828	0.9741	-3.0921	0.4258	5.9035	10.1263
14	CH_3COOCH or CH_3COOC	-0.1205	-1.5900	0.0056	5.7480	0.2637	-16.6004	-17.0885	2.5904	0.3321	-22.2984	-2.0856
15	CO-O-CO	0.0701	-1.4141	0.0000	-6.1088	-1.6082	-19.5481	-24.3719	-1.2385	-0.3014	7.6939	****
16	CHOH	-0.1849	-1.5693	0.0010	3.9998	0.1715	-13.6238	-12.9287	0.7321	0.1316	3.2443	1.0683
17	COH	-0.3371	-4.7485	0.0056	15.1556	0.6164	-26.3550	-28.3242	-0.5941	-0.0283	-11.8388	1.0698
18	CH_3COCH_nOH (n in 0..2)	-0.1226	-1.2286	-0.0052	0.0648	-0.1046	9.8524	-2.4184	****	****	****	****
19	NCCHOH or NCCOH	0.4442	0.9065	-0.0039	11.4808	-0.7565	12.5407	37.6552	0.5529	-0.4165	0.0000	****
20	OH-CH_n-COO (n in 0..2)	-0.2171	****	****	****	-1.0392	****	****	****	0.1472	-31.1800	****
21	$CH_m(OH)CH_n(OH)$ (m,n in 0..2)	0.0652	1.9395	-0.0028	13.6351	0.0219	-9.1409	-9.4404	1.1315	0.1970	11.8974	-5.6311
22	$CH_m(OH)CH_n(NH_p)$ (m,n,p in 0..2)	0.3096	4.4706	-0.0031	12.8515	-0.0778	-12.5726	-13.2008	0.6291	-0.0215	14.6940	****
23	$CH_m(NH_2)CH_n(NH_2)$ (m,n in 0..2)	0.3726	0.3128	-0.0029	15.5503	0.4916	0.4225	-4.9360	-0.9191	0.1318	38.8881	6.3363
24	$CH_m(NH)CH_n(NH_2)$ (m,n in 1..2)	0.2822	4.2200	-0.0030	****	****	1.7415	0.3214	-7.0910	0.0741	33.1648	2.4233
25	$H_2NCOCH_nCH_mCONH_2$ (m,n in 1..2)	-0.0558	****	****	****	****	****	****	****	-0.0620	****	****
26	$CH_m(NH_n)$-COOH (m,n in 0..2)	-0.0614	****	-0.0064	23.5967	1.9661	6.6270	9.8125	10.7257	-0.7893	****	****
27	HOOC-CH_n-COOH (n in 1..2)	-0.0771	19.1961	-0.0042	28.8153	1.1884	1.5448	1.4782	10.9770	0.6901	****	****
28	HOOC-CH_n-CH_m-COOH (n, m in 1..2)	-0.0345	16.2093	-0.0017	30.9875	1.9761	2.3581	0.4288	35.9435	0.3463	-8.6911	****
29	HO-CH_n-COOH (n in 1..2)	-0.1592	****	****	****	-0.2369	****	****	****	0.3755	-33.2207	****
30	NH_2-CH_n-CH_m-COOH (n, m in 1..2)	-0.1656	****	****	****	-9.0556	****	****	****	-1.4293	****	****
31	CH_3-O-CH_n-COOH (n in 1..2)	-0.2100	5.1074	-0.0030	17.8568	-1.5311	1.3721	****	-1.8939	-0.0268	****	****
32	HS-CH-COOH	-0.3486	****	****	****	-0.5790	****	****	****	0.2208	****	****
33	HS-CH_n-CH_m-COOH (n, m in 1..2)	-0.1109	5.0190	-0.0017	24.6679	1.7711	-5.2592	10.2912	0.6421	-0.0288	****	****
34	NC-CH_n-CH_m-CN (n, m in 1..2)	-0.0687	17.1845	0.0024	-9.5751	****	19.6333	18.3990	-8.4892	-0.7031	****	****
35	OH-CH_n-CH_m-CN (n, m in 1..2)	0.0308	4.9483	-0.0011	-0.3340	-1.3934	-3.1739	-1.4643	-4.4647	-0.2954	****	****

36	HS-CH_n-CH_m-SH (*n, m* in 1..2)	0.3486	6.0898	-0.0017	-17.6517	0.6309	4.3234	-0.5664	0.9008	****	****	4.0296
37	COO-CH_n-CH_m-OOC (*n, m* in 1..2)	0.0669	3.5070	-0.0023	-3.1013	-0.1619	-4.2925	-6.5362	****	0.0409	18.5312	-0.1232
38	OOC-CH_m-CH_m-COO (*n, m* in 1..2)	0.2610	4.8206	0.0009	****	0.6970	****	****	3.9216	0.0336	****	****
39	NC-CH_n-COO (*n* in 1..2)	0.3358	2.5120	0.0003	-2.8737	0.7187	-9.0684	****	0.0229	-0.4061	****	****
40	COCH_nCOO (*n* in 1..2)	0.0924	1.8790	-0.0009	0.4983	-0.3211	10.0013	-11.4766	****	0.3307	28.2659	****
41	CH_m-O-CH_n=CH_p (*m,n,p* in 0..3)	0.0567	0.6255	-0.0012	-3.6089	0.3113	-14.6024	-17.2301	1.9751	0.0031	-18.9023	6.3990
42	CH_m=CH_n-F (*m,n* in 0..2)	0.0441	1.1930	-0.0084	0.0000	0.2153	0.0000	42.5585	0.0000	0.1499	****	****
43	CH_m=CH_n-Br (*m,n* in 0..2)	-0.1808	-3.4098	-0.0110	4.4217	-0.4046	12.5062	12.0983	1.7988	0.2119	****	****
44	CH_m=CH_n-I (*m,n* in 0..2)	-0.3108	****	****	****	0.2906	****	****	****	****	****	****
45	CH_m=CH_n-Cl (*m,n* in 0..2)	-0.0239	-0.1427	-0.0025	0.4361	0.1547	6.5775	-0.9991	-1.6875	0.3476	20.6823	3.3906
46	CH_m=CH_n-CN (*m,n* in 0..2)	-0.1070	-0.5768	0.0071	39.8619	0.0840	-11.8012	-46.9867	-0.6471	-0.1624	-15.2287	3.1620
47	CH_n=CH_m-COO-CH_p (*m,n,p* in 0..3)	0.0861	-0.7358	-0.0039	1.9955	0.3601	-35.1022	-40.6102	-0.6289	0.0150	-0.9384	****
48	CH_m=CH_n-CHO (*m,n* in 0..2)	0.2191	1.7653	0.0089	****	-0.5134	****	****	****	0.3320	-2.1667	****
49	CH_m=CH_n-COOH (*m,n* in 0..2)	0.1364	-1.1866	-0.0032	14.1051	1.6148	10.0524	2.0204	5.6921	0.7099	27.4890	****
50	aC-CH_n-X (*n* in 1..2) X: Halogen	0.2251	7.6911	0.0018	-15.0334	0.4797	-5.2913	-5.0835	1.5124	0.0519	31.5019	****
51	aC-CH_n-NH_m (*n* in 1..2; *m* in 0..2))	-0.0423	2.8668	0.0088	-2.0477	-0.3004	-4.6214	-13.4697	-1.2624	-0.4910	****	2.6501
52	aC-CH_n-O- (*n* in 1..2)	0.0914	0.0799	0.0002	-15.2270	0.6859	3.2914	4.3859	1.1046	-0.1893	0.0000	****
53	aC-CH_n-OH (*n* in 1..2)	0.0714	-0.6470	0.0003	-11.5271	0.7467	-4.2548	-1.3131	0.1322	-0.0420	4.2772	****
54	aC-CH_n-CN (*n* in 1..2)	0.1374	10.8883	0.0071	16.1808	0.2712	24.3740	-24.3284	****	-1.0253	****	****
55	aC-CH_n-CHO (*n* in 1..2)	0.2908	****	****	****	2.4059	****	****	****	0.3198	****	****
56	aC-CH_n-SH (*n* in 1..2)	0.1594	6.4747	0.0105	-19.7643	-0.2365	14.1436	****	0.4251	****	****	****
57	aC-CH_n-COOH (*n* in 1..2)	0.2050	6.4527	-0.0012	****	1.6724	****	****	3.7099	0.2326	****	****
58	aC-CH_n-CO- (*n* in 1..2)	0.1651	****	****	****	0.2774	****	****	****	-0.3167	****	****
59	aC-CH_n-S- (*n* in 1..2)	-0.1400	****	****	****	0.5177	****	****	****	-0.5420	****	****
60	aC-CH_n-OOC-H (*n* in 1..2)	0.2505	4.3834	0.0043	-12.1801	1.1002	2.3398	4.1991	****	****	****	****
61	aC-CH_n-NO_2 (*n* in 1..2)	0.2936	****	****	****	0.1159	****	-23.4033	****	-0.4890	****	****
62	aC-CH_n-$CONH_2$ (*n* in 1..2)	-0.1632	****	****	****	1.8575	****	****	****	-0.4038	****	****
63	aC-CH_n-OOC (*n* in 1..2)	-0.0331	3.2086	-0.0004	-14.7528	-0.0831	-3.5017	0.0225	****	0.1382	25.1873	****
64	aC-CH_n-COO (*n* in 1..2)	0.1318	****	****	****	0.6322	****	****	****	-0.3490	****	****
65	aC-SO_2-OH	-0.1235	****	****	0.0000	0.7846	0.0000	****	****	-1.0081	****	****
66	aC-C$H(CH_3)_2$	0.1303	0.3642	-0.0043	3.7173	0.3193	9.2647	5.4369	-0.6901	0.0603	6.2832	1.9723
67	aC-C$(CH_3)_3$	0.1686	8.3074	-0.0048	-3.8927	0.2962	-2.4071	5.5887	1.0386	0.1255	0.0000	-1.8302
68	aC-CF_3	-0.1097	-4.5292	****	****	1.0015	****	****	****	-0.4160	****	2.9422
69	$(CH_n$=C$)_{cyc}$-CHO (*n* in 0..2)	0.1985	6.8126	-0.0026	3.6063	-0.6843	-6.2902	-0.7938	-1.9821	0.4771	19.4540	****
70	$(CH_n$=C$)_{cyc}$-COO-CH_m (*n,m* in 0..3)	0.1548	****	****	****	-0.0664	****	****	****	-0.0224	****	****
71	$(CH_n$=C$)_{cyc}$-CO- (*n* in 0..2)	0.4200	****	****	****	0.7333	****	****	****	0.3526	****	****
72	$(CH_n$=C$)_{cyc}$-CH_3 (*n* in 0..2)	-0.0378	-0.5343	0.0024	-1.9812	-0.2646	17.5547	12.7329	0.8845	-0.0363	-20.0682	0.5401
73	$(CH_n$=C$)_{cyc}$-CH_2 (*n* in 0..2)	-0.0942	-3.1684	0.0027	-7.9348	-0.8257	-24.0807	****	-0.1116	-0.2892	-2.0815	****
74	$(CH_n$=C$)_{cyc}$-CN (*n* in 0..2)	-0.3094	****	****	****	2.6108	****	****	****	-0.0471	****	-1.1442

75	$(CH_n{=}C)_{cyc}$-Cl (*n* in 0..2)	-0.0988	****	****	****	-0.4751	****	****	****	0.3371	****	****
76	CH_{cyc}-CH_3	-0.1675	-0.3458	0.0033	-1.7454	0.0997	18.4415	86.8776	0.2756	-0.0719	-13.2791	0.2465
77	CH_{cyc}-CH_2	-0.1303	0.3439	0.0033	0.1589	-0.0436	8.6015	78.6531	-0.2009	-0.0167	5.8900	-0.2872
78	CH_{cyc}-CH	0.0319	0.6262	0.0009	-2.9086	0.5716	10.7621	80.5539	3.4484	0.3554	12.2737	1.0578
79	CH_{cyc}-C	-0.0642	3.7347	****	****	0.8971	****	****	****	0.1405	****	1.9705
80	CH_{cyc}-CH=CH_n (*n* in 1..2)	-0.1998	-0.2938	0.0014	-11.2524	-0.1997	12.9333	80.9884	****	0.5718	8.0519	5.9104
81	CH_{cyc}-C=CH_n (*n* in 1..2)	0.0223	****	****	****	-0.4540	****	****	****	0.4178	19.6550	****
82	CH_{cyc}-Cl	-0.0499	-3.0141	****	****	0.9872	****	****	****	-0.0383	****	****
83	CH_{cyc}-F	-0.0940	****	****	****	3.9892	****	****	****	-0.3161	****	****
84	CH_{cyc}-OH	-0.3381	2.7089	0.0045	****	0.4001	-1.0379	62.9031	-6.3662	0.3092	-6.7951	3.9273
85	CH_{cyc}-NH_2	-0.3556	-2.8873	0.0095	****	-0.0731	5.3863	25.2271	0.5535	-0.0570	-22.5848	-8.4785
86	CH_{cyc}-NH-CH_n (*n* in 0..3)	-0.2411	-2.8390	-0.0017	****	-0.3622	-21.1409	48.2957	3.7041	-0.0615	****	****
87	CH_{cyc}-N-CH_n (*n* in 0..3)	-0.0081	****	****	****	0.6709	****	****	****	0.1113	****	****
88	CH_{cyc}-SH	-0.1619	2.5523	0.0109	****	****	-13.3787	58.5064	-0.4609	****	****	0.0000
89	CH_{cyc}-CN	0.3281	****	****	****	0.7858	****	****	****	-0.7459	****	-3.9451
90	CH_{cyc}-COOH	0.1814	****	****	****	0.1389	****	****	****	0.4730	****	****
91	CH_{cyc}-CO	0.0176	****	****	****	0.3761	****	****	****	-0.2120	****	3.2566
92	CH_{cyc}-NO_2	-0.2571	****	****	****	-0.6048	****	67.4364	****	-0.0908	****	****
93	CH_{cyc}-S-	-0.1065	****	****	****	-0.5508	****	****	****	-0.4044	****	****
94	CH_{cyc}-CHO	-0.0915	****	****	****	0.3613	****	****	****	****	****	****
95	CH_{cyc}-O-	-0.0402	****	****	****	0.0483	****	****	****	-0.2341	****	****
96	CH_{cyc}-OOCH	-0.2720	****	****	****	****	****	****	****	0.2393	****	****
97	CH_{cyc}-COO	0.1126	****	****	****	0.4482	****	****	****	-0.7404	****	0.0000
98	CH_{cyc}-OOC	-0.3484	-0.4978	0.0000	****	0.4256	5.2261	53.8488	****	0.0744	1.2273	****
99	C_{cyc}-CH_3	0.0019	0.1796	0.0017	1.9275	-0.0802	2.9003	-23.3439	-3.0371	0.0704	-0.7406	0.7944
100	C_{cyc}-CH_2	0.1739	0.5885	-0.0001	-7.5226	-0.6108	12.6311	-26.0667	****	0.0862	****	1.5449
101	C_{cyc}-OH	-0.3669	-4.0786	-0.0035	-13.8695	0.8454	-26.0491	-64.7668	-0.7048	0.3348	-10.5522	****
102	$>N_{cyc}$-CH_3	-0.0808	-0.0001	0.0057	0.0000	0.1915	0.0000	****	0.0000	-0.1446	0.0000	****
103	$>N_{cyc}$-CH_2	-0.1738	****	****	****	-0.5876	****	-2.6474	****	-0.2832	****	****
104	AROMRINGs^{1}s^2	-0.0035	-0.5587	0.0028	12.8292	-0.6360	2.7645	3.0600	-2.9135	0.0092	-24.2331	1.6730
105	AROMRINGs^{1}s^3	0.0506	-1.2269	0.0030	15.8491	-0.6253	0.1061	3.4173	-2.9614	0.1582	-3.4558	-2.2627
106	AROMRINGs^{1}s^4	0.0929	-0.5134	0.0022	12.8037	0.1571	1.6039	3.1108	-0.1208	0.0730	-11.3047	1.1951
107	AROMRINGs^{1}s^{2}s^3	0.0577	1.0648	-0.0026	-11.4318	-0.0456	8.1841	6.1402	-2.0809	-0.2035	-25.6460	4.1320
108	AROMRINGs^{1}s^{2}s^4	0.0631	-0.4645	-0.0015	-12.2999	-0.4116	2.4368	1.1269	-2.1164	0.0869	-26.5119	4.7221
109	AROMRINGs^{1}s^{3}s^5	0.0175	-1.4995	0.0088	-15.5715	0.4333	1.9842	-5.8803	-2.4732	0.3220	-1.7193	6.7070
110	AROMRINGs^{1}s^{2}s^{3}s^4	0.1411	5.9314	-0.0065	-29.5053	-0.2450	19.5483	20.7707	-2.5691	-0.0693	-19.8200	****
111	AROMRINGs^{1}s^{2}s^{3}s^5	0.0772	2.5689	-0.0024	-22.5053	-0.2270	10.2483	15.5257	-3.0791	0.0009	-24.4870	****
112	AROMRINGs^{1}s^{2}s^{4}s^5	0.1318	1.6886	-0.0024	-8.1292	0.5634	4.8483	12.1457	7.2009	0.2543	****	****
113	PYRIDINEs^2	0.0425	-2.1157	0.0079	0.7542	-0.4206	-12.0456	-5.5953	-3.1210	-0.4911	-5.7548	-3.1446

114	PYRIDINEs^3	0.1362	2.1936	0.0079	0.7542	-0.4951	-4.8556	0.6709	1.3390	-0.7522	****	-1.0746
115	PYRIDINEs^4	0.2174	2.4148	0.0032	6.3542	0.2343	-7.0356	-1.5652	36.4353	-0.6147	12.3552	-0.8246
116	PYRIDINEs^{2}s^3	0.1462	0.0573	****	-16.8254	-0.0667	****	-2.5750	****	-0.1985	****	-0.7951
117	PYRIDINEs^{2}s^4	0.1993	-1.6290	0.0051	-16.8254	3.9162	****	-5.9543	****	0.0893	****	-1.0451
118	PYRIDINEs^{2}s^5	0.0780	-2.1484	0.0071	-16.8254	0.4767	****	-0.5351	****	-0.4426	****	-0.7051
119	PYRIDINEs^{2}s^6	0.1293	-5.8586	0.0086	****	0.0109	-16.9493	-10.1123	-4.4058	-0.1165	****	-3.1651
120	PYRIDINEs^{3}s^4	0.5686	6.2514	0.0003	-16.8254	0.2094	****	-0.8972	****	-0.4759	****	2.0049
121	PYRIDINEs^{3}s^5	0.1820	2.6197	****	****	0.2108	****	1.5996	****	-0.1402	****	0.9549
122	PYRIDINEs^{2}s^{3}s^6	-0.0605	****	****	****	****	****	****	****	****	****	****
123	(CH^n=CH^m)^{cyc}-COOH	0.2498	****	****	****	2.8876	****	****	****	0.4535	****	****
124	AROMRINGs^{1}s^{2}s^{3}s^{4}s^5	0.0605	****	****	****	-0.4199	****	****	-4.7340	0.0131	****	0.1904
125	aC-NHCOCH$_2$N	-0.1402	****	****	****	-5.3116	****	****	****	-0.2844	****	****
126	(N=C)$_{cyc}$-CH$_3$	0.1498	****	****	****	2.6716	****	48.2349	****	-0.3653	****	****
127	aC-CONH(CH$_2$)$_2$N	-0.0813	****	****	****	-2.4284	****	****	****	-0.2128	****	****
128	aC-SO$_2$NH$_n$ (n>=0;n<3)	-0.0657	****	****	****	0.6548	****	****	****	0.4146	****	****
129	aC-SO$_2$NH$_n$ (n>=0;n<3)	-0.0002	****	****	****	0.0725	****	****	****	0.3093	****	****
130	aC-SO$_2$NH$_n$ (n>=0;n<3)	-0.0326	****	****	****	-0.9750	****	****	****	0.6492	****	****

[a] The symbols T_{b2j}, T_{c2j}, P_{c2j}, V_{c2j}, T_{m2j}, $\Delta_f G^o{}_{2j}$, $\Delta_f H^o{}_{gas2j}$, $\Delta_{fus}H_{2j}$, $LogK_{ow2j}$, F_{p2j}, and $\Delta_{vap}H^o{}_{2j}$ represent the contributions (D_j) of the second-order groups for the corresponding properties

Table A6 (Continued). MG method based property models analysed using simultaneous regression method: Second-order groups and their contributions [a] for the properties÷ $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, AiT, ω and V_m

	Group	$\Delta_{vap}H_{2j}$	$\Delta_{vap}S_{2j}$	δ_{D2j}	δ_{P2j}	δ_{H2j}	δ_{2j}	$AiT1a_{2j}$	$AiT2b_{2j}$	ω_{2j}	V_{m2j}
1	$(CH_3)_2CH$	-0.4561	-0.0932	-0.2581	0.2698	-0.1884	-0.3119	0.0213	30.9432	-0.0017	0.0021
2	$(CH_3)_3C$	-0.4803	0.8213	0.1777	-0.0817	2.3617	0.2229	0.1323	69.4951	-0.0010	0.0045
3	$CH(CH_3)CH(CH_3)$	0.7318	0.3800	0.0518	-0.5663	-0.0999	0.4113	-0.0300	39.6769	-0.0009	-0.0019
4	$CH(CH_3)C(CH_3)_2$	0.7690	-0.2184	****	****	****	1.1140	-0.1255	6.2775	-0.0012	-0.0034
5	$C(CH_3)_2C(CH_3)_2$	****	****	****	****	****	1.1353	-0.0690	-31.2791	-0.0008	-0.0066
6	CH_n=CH_m-CH_p=CH_k (k,m,n,p in 0..2)	0.8048	-0.6273	-0.2768	-3.1365	-1.5359	0.0348	-0.0247	-51.9915	-0.0008	-0.0014
7	CH_3-CH_m=CH_n (m,n in 0..2)	-0.1269	0.0588	-0.3081	-1.5300	-1.0401	-0.2313	0.0613	52.1765	-0.0006	0.0027
8	CH_2-CH_m=CH_n (m,n in 0..2)	0.9044	1.0654	-0.2360	-1.6367	-0.5197	-0.1940	0.1237	61.7444	0.0002	0.0012
9	CH_p-CH_m=CH_n (m,n in 0..2; p in 0..1)	****	****	-0.2796	0.8006	-0.7440	-0.4152	10.0739	164.2867	-0.0022	0.0020
10	CHCHO or CCHO	****	****	-0.4596	-2.3384	0.8469	-1.3461	-0.3487	-36.7985	0.0005	0.0021
11	CH_3COCH_2	0.4522	-1.5937	0.0530	-1.2685	-0.3451	-0.5303	-0.5253	-91.9394	-0.0010	0.0003
12	CH_3COCH or CH_3COC	1.8853	1.4706	-0.2455	-2.1456	-0.5600	-0.0954	****	****	-0.0003	0.0003
13	CHCOOH or CCOOH	****	****	0.1455	1.6507	2.4286	0.9480	4.8993	51.0637	-0.0016	0.0050
14	CH_3COOCH or CH_3COOC	-2.2837	-1.7525	-0.2666	0.6136	0.1687	-0.1911	-0.3815	-133.9296	-0.0024	0.0023
15	CO-O-CO	0.6666	0.5120	-0.1879	2.5450	1.4428	0.5379	0.4023	-134.7961	0.0022	0.0024
16	CHOH	-2.6237	-3.2560	0.0394	0.1176	0.6216	-0.4165	-0.2087	-46.5762	0.0004	0.0001
17	COH	-2.7937	0.1906	-0.0212	-1.0511	-0.8848	-0.1556	-0.4370	-294.3050	0.0025	0.0038
18	CH_3COCH_nOH (n in 0..2)	****	****	****	****	****	1.6911	****	****	0.0072	-0.0026
19	NCCHOH or NCCOH	****	****	1.1107	3.2379	-1.1972	2.0962	-0.3009	-19.7984	0.0062	0.0030
20	OH-CH_n-COO (n in 0..2)	****	****	-0.1376	-0.1398	-3.3827	****	0.0992	-32.9962	0.0047	0.0012
21	$CH_m(OH)CH_n(OH)$ (m,n in 0..2)	-5.6244	-13.4425	0.1682	-1.9454	-3.2166	-0.6207	-0.2607	-75.2269	-0.0036	0.0020
22	$CH_m(OH)CH_n(NH_p)$ (m,n,p in 0..2)	1.3129	-4.5502	0.6644	1.5311	-0.8175	2.0921	-0.5851	-176.0792	-0.0003	0.0009
23	$CH_m(NH_2)CH_n(NH_2)$ (m,n in 0..2)	3.5002	1.9697	0.2010	-1.6202	3.4033	2.6346	-0.2902	-50.8837	-0.0020	-0.0009
24	$CH_m(NH)CH_n(NH_2)$ (m,n in 1..2)	1.1028	-1.7255	0.1405	1.8810	-0.2500	1.0660	-0.2641	-42.3587	0.0007	0.0057
25	$H_2NCOCH_nCH_mCONH_2$ (m,n in 1..2)	****	****	****	****	****	****	****	****	****	****
26	$CH_m(NH_n)$-COOH (m,n in 0..2)	****	****	****	****	****	-4.0094	****	****	-0.0043	0.0155
27	HOOC-CH_n-COOH (n in 1..2)	****	****	****	****	****	6.3081	****	****	0.0002	-0.0058
28	HOOC-CH_n-CH_m-COOH (n, m in 1..2)	****	****	****	****	****	-3.0775	-5.0329	-256.1538	0.0012	0.0012
29	HO-CH_n-COOH (n in 1..2)	****	****	0.4733	1.0034	9.4799	3.5131	8.2115	21.6037	0.0070	0.0171
30	NH_2-CH_n-CH_m-COOH (n, m in 1..2)	****	****	****	****	****	****	****	****	****	****
31	CH_3-O-CH_n-COOH (n in 1..2)	****	****	****	****	****	5.3782	****	****	-0.0030	-0.0063
32	HS-CH-COOH	****	****	-0.9666	1.6226	2.9668	4.7479	****	****	-0.0009	-0.0042
33	HS-CH_n-CH_m-COOH (n, m in 1..2)	****	****	****	****	****	2.3141	****	****	-0.0013	-0.0032
34	NC-CH_n-CH_m-CN (n, m in 1..2)	4.9761	0.5426	0.7714	-5.4527	1.2713	0.1776	****	****	-0.0007	-0.0010
35	OH-CH_n-CH_m-CN (n, m in 1..2)	****	****	0.5877	2.9021	2.7001	1.8320	0.0948	-12.0092	0.0036	-0.0030

36	HS-CH_n-CH_m-SH (*n, m* in 1..2)	2.9250	1.7209	0.8012	-0.4747	0.2170	2.1166	****	****	0.0004	0.0001
37	COO-CH_n-CH_m-OOC (*n, m* in 1..2)	****	****	-0.2289	2.4076	0.4053	0.1617	0.2323	****	0.0000	0.0000
38	OOC-CH_n-CH_m-COO (*n, m* in 1..2)	****	****	-0.9622	3.5347	1.1963	0.2834	****	****	0.0021	-0.0012
39	NC-CH_n-COO (*n* in 1..2)	6.0367	8.1500	0.1665	-3.9012	0.6450	1.1237	****	****	-0.0067	-0.0012
40	$COCH_nCOO$ (*n* in 1..2)	****	****	0.0971	0.6223	0.4474	0.6686	10.6032	-46.6300	0.0026	-0.0016
41	CH_m-O-CH_n=CH_p (*m,n,p* in 0..3)	0.6374	-1.2320	-0.1456	-1.0741	0.2445	0.2080	0.0337	-5.2175	-0.0008	0.0005
42	CH_m=CH_n-F (*m,n* in 0..2)	****	****	0.8039	-3.5959	-4.4333	-0.8201	0.2419	71.8329	-0.0005	0.0069
43	CH_m=CH_n-Br (*m,n* in 0..2)	****	****	-0.8204	-0.7914	-0.6056	-0.3031	****	****	0.0017	0.0002
44	CH_m=CH_n-I (*m,n* in 0..2)	****	****	0.1490	-2.1394	0.4795	****	****	****	****	****
45	CH_m=CH_n-Cl (*m,n* in 0..2)	-2.5398	-2.1866	0.2520	0.0461	-0.7345	-0.4285	0.1172	97.9028	-0.0008	0.0027
46	CH_m=CH_n-CN (*m,n* in 0..2)	0.4335	4.4662	-0.1696	-1.3468	1.6540	1.1350	-0.0277	33.6184	0.0023	0.0012
47	CH_n=CH_m-COO-CH_p (*m,n,p* in 0..3)	1.6243	3.0045	-0.1152	-1.9466	-0.8720	-0.8272	-0.3638	-178.7648	-0.0058	0.0019
48	CH_m=CH_n-CHO (*m,n* in 0..2)	0.7518	-2.3128	0.4924	-0.1767	-0.3584	0.5079	41.3681	107.5145	-0.0016	-0.0018
49	CH_m=CH_n-COOH (*m,n* in 0..2)	****	****	0.4276	-0.4406	2.1610	0.6972	-0.6595	-346.6302	-0.0015	0.0046
50	aC-CH_n-X (*n* in 1..2) X: Halogen	****	****	0.6074	0.6070	0.6689	-0.1102	0.1165	53.6608	-0.0009	0.0001
51	aC-CH_n-NH_m (*n* in 1..2; *m* in 0..2))	****	****	-0.4194	-2.1514	1.9559	-3.5506	****	****	-0.0012	0.0013
52	aC-CH_n-O- (*n* in 1..2)	****	****	-0.4615	0.1897	0.8380	0.0121	0.5613	-8.4028	0.0009	-0.0013
53	aC-CH_n-OH (*n* in 1..2)	****	****	0.4381	0.3564	0.8857	-0.1387	4.9890	259.6343	-0.0004	0.0010
54	aC-CH_n-CN (*n* in 1..2)	****	****	0.8337	0.6377	1.9122	0.1165	****	****	0.0019	0.0013
55	aC-CH_n-CHO (*n* in 1..2)	****	****	****	****	****	****	****	****	0.0004	-0.0009
56	aC-CH_n-SH (*n* in 1..2)	****	****	****	****	****	-0.5739	****	****	-0.0012	0.0032
57	aC-CH_n-COOH (*n* in 1..2)	****	****	****	****	****	****	****	****	0.0052	0.0271
58	aC-CH_n-CO- (*n* in 1..2)	****	****	-1.4637	****	****	****	****	****	****	****
59	aC-CH_n-S- (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****
60	aC-CH_n-OOC-H (*n* in 1..2)	****	****	****	****	****	0.3459	0.4344	113.7271	-0.0016	0.0008
61	aC-CH_m-NO_2 (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****
62	aC-CH_n-$CONH_2$ (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****
63	aC-CH_n-OOC (*n* in 1..2)	-1.9711	-2.5093	-1.4361	1.6473	0.0979	-0.4009	****	****	-0.0005	0.0064
64	aC-CH_n-COO (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****
65	aC-SO_2-OH	****	****	****	****	****	****	****	****	****	****
66	aC-$CH(CH_3)_2$	****	****	1.3456	-1.5156	1.8781	1.6291	0.0863	67.1581	0.0007	0.0014
67	aC-$C(CH_3)_3$	****	****	0.0656	-0.5074	1.9725	2.1807	0.1842	26.0844	-0.0007	0.0019
68	aC-CF_3	0.0000	0.0000	4.6154	-6.8852	-8.8009	-0.2885	****	****	-0.0009	0.0004
69	$(CH_n{=}C)_{(cyc)}$-CHO (*n* in 0..2)	3.8300	1.7993	0.6376	-0.2193	-3.5531	0.2112	-0.4336	21.3102	-0.0059	0.0221
70	$(CH_n{=}C)_{cyc}$-COO-CH_m (*n,m* in 0..3)	****	****	-0.7416	-4.8547	1.5749	****	****	****	****	****
71	$(CH_n{=}C)_{cyc}$-CO- (*n* in 0..2)	****	****	-0.5027	-1.2422	2.4173	0.1720	****	****	****	****
72	$(CH_n{=}C)_{cyc}$-CH_3 (*n* in 0..2)	-3.6186	-3.9575	0.6248	-2.9668	1.5925	-1.1044	-0.1405	-20.7562	-0.0034	0.0221
73	$(CH_n{=}C)_{cyc}$-CH_2 (*n* in 0..2)	3.3394	6.6328	-1.0210	-5.5317	-0.1157	1.0543	-0.0913	-41.0791	-0.0014	0.0216
74	$(CH_n{=}C)_{cyc}$-CN (*n* in 0..2)	****	****	0.1248	-1.8783	3.0012	****	****	****	****	****

75	$(CH_n{=}C)_{cyc}$-Cl (*n* in 0..2)	****	****	0.8143	-4.3334	2.7380	1.3809	****	****	-0.0051	-0.0079
76	CH_{cyc}-CH_3	-0.3912	-0.3375	-1.1526	1.0660	-2.6027	0.8659	0.3858	56.5571	0.0095	0.0029
77	CH_{cyc}-CH_2	-0.2615	-1.6235	-0.2448	-0.5983	-4.0908	0.0197	0.1973	10.9994	0.0109	0.0019
78	CH_{cyc}-CH	0.8542	0.0960	-3.4025	-2.1028	-1.5266	0.3056	0.0671	-16.2415	0.0104	-0.0062
79	CH_{cyc}-C	0.6014	-0.2373	****	****	****	****	****	****	0.0123	0.0056
80	CH_{cyc}-$CH{=}CH_n$ (*n* in 1..2)	0.5881	-2.6668	-0.6984	-0.5937	-3.3474	2.0400	0.6095	73.2595	0.0097	0.0064
81	CH_{cyc}-$C{=}CH_n$ (*n* in 1..2)	****	****	-1.6598	-2.2489	-1.5637	2.9189	1.1049	62.1818	0.0153	-0.0028
82	CH_{cyc}-Cl	****	****	0.5687	4.4650	-1.2717	0.6430	****	****	****	****
83	CH_{cyc}-F	****	****	****	****	****	****	****	****	****	****
84	CH_{cyc}-OH	2.3625	10.2832	-0.1237	2.2395	-1.3628	-2.2368	-0.2600	-108.4679	0.0105	-0.0007
85	CH_{cyc}-NH_2	-4.5889	-5.2424	-1.5497	-2.6012	-5.1402	-4.5687	-0.5278	-47.7291	0.0102	0.0036
86	CH_{cyc}-NH-CH_n (*n* in 0..3)	****	****	****	****	****	-0.2795	****	****	0.0100	0.0027
87	CH_{cyc}-N-CH_n (*n* in 0..3)	****	****	****	****	****	****	****	****	****	****
88	CH_{cyc}-SH	0.0000	0.0000	****	****	****	-0.6601	****	****	0.0100	0.0048
89	CH_{cyc}-CN	0.6518	1.5388	8.7979	8.7539	3.0634	****	****	****	0.0118	0.0061
90	CH_{cyc}-COOH	****	****	****	****	****	0.1992	****	****	0.0098	0.0050
91	CH_{cyc}-CO	3.0081	1.4526	7.2411	6.9033	11.8459	2.1873	****	****	0.0157	-0.0067
92	CH_{cyc}-NO_2	****	****	****	****	****	****	****	****	****	****
93	CH_{cyc}-S-	****	****	****	****	****	****	****	****	0.0084	0.0031
94	CH_{cyc}-CHO	****	****	0.1109	0.9089	-3.5562	0.4012	****	****	0.0088	0.0048
95	CH_{cyc}-O-	****	****	-0.4245	1.4219	-1.0584	-1.4860	****	****	0.0110	0.0062
96	CH_{cyc}-OOCH	****	****	****	****	****	1.4056	****	****	0.0088	0.0019
97	CH_{cyc}-COO	0.1036	-0.4499	-1.2560	0.5875	-1.1121	****	****	****	0.0031	0.0031
98	CH_{cyc}-OOC	****	****	1.2601	7.9886	-0.0407	1.5543	-0.0212	-65.3315	0.0067	-0.0368
99	C_{cyc}-CH_3	-0.2911	-0.2763	-1.8638	-1.9505	-0.3430	-0.2177	1.1582	46.1762	-0.0002	-0.0682
100	C_{cyc}-CH_2	-1.0374	-2.9501	****	****	****	-4.3488	-0.7823	-145.9957	0.0015	-0.0722
101	C_{cyc}-OH	****	****	0.1917	0.8769	-1.6146	-0.9897	-0.2133	-108.5533	-0.0060	-0.0719
102	$>N_{cyc}$-CH_3	****	****	1.2465	1.7133	1.9504	0.3829	-0.1078	9.1670	0.0012	-0.0049
103	$>N_{cyc}$-CH_2	****	****	0.6026	-0.5161	-0.2783	-0.3566	****	****	0.0052	0.0042
104	AROMRINGs^{1}s^2	1.8394	3.7003	-0.2022	2.1148	-0.1562	0.1562	0.1684	33.4379	-0.0004	-0.0069
105	AROMRINGs^{1}s^3	1.3105	2.7400	-0.2887	3.3357	-0.0298	0.6763	0.2196	122.8660	0.0002	-0.0049
106	AROMRINGs^{1}s^4	1.7438	3.2160	-0.1469	1.5376	-0.4144	-0.0806	0.2044	110.8972	0.0005	-0.0015
107	AROMRINGs^{1}s^{2}s^3	****	****	0.2024	3.6173	-0.5336	-0.1551	8.2085	921.5129	0.0011	-0.0035
108	AROMRINGs^{1}s^{2}s^4	-0.1730	1.2091	-0.0667	2.4882	-0.5098	0.8872	0.2780	24.5010	0.0009	-0.0048
109	AROMRINGs^{1}s^{3}s^5	****	****	-0.3769	2.2296	-0.1894	-0.5621	5.8941	935.5989	0.0002	-0.0036
110	AROMRINGs^{1}s^{2}s^{3}s^4	****	****	0.2026	1.3547	-0.6207	0.6205	0.5482	204.6875	0.0014	-0.0118
111	AROMRINGs^{1}s^{2}s^{3}s^5	****	****	-0.5850	-0.3696	-2.0195	-3.0447	0.5484	205.1828	0.0006	-0.0085
112	AROMRINGs^{1}s^{2}s^{4}s^5	****	****	-0.0562	3.3194	0.4745	-2.3553	0.5485	205.2480	0.0013	-0.0051
113	PYRIDINEs^2	-1.5196	-3.6456	-0.3903	-2.3308	-0.7000	0.9455	0.0082	47.7136	0.0018	-0.0002

114	PYRIDINEs^{3}	-0.3396	-3.9827	****	****	****	0.0636	0.0457	54.1609	0.0008	-0.0015
115	PYRIDINEs^{4}	-0.1796	-3.8570	-0.3062	2.3041	-0.8632	1.5037	0.0090	47.8203	0.0019	-0.0012
116	PYRIDINEs^{2}s^{3}	-0.2193	-2.4358	****	****	****	****	****	****	****	****
117	PYRIDINEs^{2}s^{4}	-0.7693	-3.1265	****	****	****	****	****	****	****	****
118	PYRIDINEs^{2}s^{5}	-0.6193	-2.6958	****	****	****	2.0753	****	****	****	****
119	PYRIDINEs^{2}s^{6}	-1.8393	-2.6314	****	****	****	0.6654	****	****	0.0024	-0.0008
120	PYRIDINEs^{3}s^{4}	0.6907	-3.9842	****	****	****	****	****	****	****	****
121	PYRIDINEs^{3}s^{5}	0.1607	-3.7447	****	****	****	****	****	****	****	****
122	PYRIDINEs^{2}s^{3}s^{6}	****	****	****	****	****	****	****	****	****	****
123	(CH$''$=CH$'''$)^{cyc}-COOH	****	****	****	****	****	****	****	****	****	****
124	AROMRINGs^{1}s^{2}s^{3}s^{4}s^{5}	0.4787	0.7310	****	****	****	****	****	****	0.0007	-0.0148
125	aC-NHCOCH$_2$N	****	****	****	****	****	****	****	****	****	****
126	(N=C)$_{cyc}$-CH$_3$	****	****	****	****	****	****	****	****	****	****
127	aC-CONH(CH$_2$)$_2$N	****	****	****	****	****	****	****	****	****	****
128	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	****	****	****	****	****	****	****	****
129	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	****	****	****	****	****	****	****	****
130	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	-2.4807	-3.0742	-7.0834	****	****	****	****	****

[a] The symbols $\Delta_{vap}H_{2j}$, $\Delta_{vap}S_{2j}$, δ_{D2j}, δ_{P2j}, δ_{H2j}, Ait_{a2j}, Ait_{b2j}, ω_{2j}, and V_{m2j} represent the contributions (D_j) of the second-order groups for the corresponding properties.

Table A7. MG method based property models analysed using simultaneous regression method: Third-order groups and their contributions [a] for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, and $\Delta_{vap}H^o$

	Group	T_{b3k}	T_{c3k}	P_{c3k}	V_{c3k}	T_{m3k}	$\Delta_f G^o$	$\Delta_f H^o_{gas3k}$	$\Delta_{fus}H_{3k}$	$LogK_{ow3k}$	F_{p3k}	$\Delta_{vap}H^o{}_{3k}$
1	HOOC-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	13.4712	-0.0009	11.0007	0.6004	-0.0255	0.4300	-0.0476	-0.0374	-16.1958	****
2	NH_n-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	-2.3828	****	****
3	NH_2-$(CH_n)_m$-OH (m>2, n in 0..2)	****	5.5882	-0.0004	-24.5991	0.4112	0.0278	-0.6488	-0.0176	0.0077	****	****
4	OH-$(CH_n)_m$-OH (m>2, n in 0..2)	****	4.1384	-0.0021	-3.9094	-0.3668	-0.1181	1.2575	0.0370	-0.6217	30.4538	0.7692
5	OH-$(CH_p)_k$-O-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****	****	****
6	OH-$(CH_p)_k$-S-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****	****	****
7	OH-$(CH_p)_k$-NH_x-$(CH_n)_m$-OH (m,k>0; p,n,x in 0..2)	****	****	****	****	****	****	****	****	****	****	****
8	CH_p-O-$(CH_n)_m$-OH (m>2; n,p in 0..2)	****	****	****	****	****	****	****	****	****	****	****
9	NH_2-$(CH_n)_m$-NH_2 (m>2; n in 0..2))	0.1929	-1.8632	0.0057	-0.7610	0.3867	-0.0303	-1.7688	0.0956	0.4177	****	6.7152
10	NH_k-$(CH_n)_m$-NH_2 (m>2; k in 0..1; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****
11	SH-$(CH_n)_m$-SH (m>2; n in 0..2)	0.4409	****	****	****	-0.6229	****	****	****	****	****	4.4309
12	NC-$(CH_n)_m$-CN (m>2)	****	16.6171	0.0020	-15.8182	-0.1407	-0.0367	-0.5751	0.1697	-0.9615	-3.3177	****
13	COO-$(CH_n)_m$-OOC (m>2; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****
14	aC-$(CH_n$=$CH_m)_{cyc}$ (fused rings) (n,m in 0..1)	-0.3645	-9.2419	0.0121	-3.8125	-0.1106	-0.0283	8.1046	-0.0418	0.1020	****	****
15	aC-aC (different rings)	-0.0067	-2.3599	-0.0020	28.9750	-0.2793	0.1102	4.4765	0.1048	0.3278	-17.3935	****
16	aC-CH_{ncyc} (different rings) (n in 0..1)	-0.2718	-2.7789	0.0070	****	-0.2608	-0.0565	72.4658	0.0097	-0.0494	****	0.0000
17	aC-CH_{ncyc} (fused rings) (n in 0..1)	-0.3564	-6.1429	0.0092	-5.1805	-0.3781	0.0243	12.6127	0.0149	0.0214	****	-1.2373
18	aC-$(CH_n)_m$-aC (different rings) (m>1; n in 0..2)	0.5407	11.8299	0.0010	-6.0599	0.1356	0.0406	2.2023	12.6038	0.1092	15.4464	****
19	aC-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	-0.2127	****	****	****	-1.0198	****	****	****	-0.2033	****	****
20	CH_{cyc}-CH_{cyc} (different rings)	-0.1042	6.2025	0.0018	****	1.3317	****	****	****	0.2971	2.3374	****
21	CH_{cyc}-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****
22	$CH_{multiring}$	0.1245	1.9976	-0.0015	4.3269	0.6170	0.0478	60.1441	-0.0023	-0.2261	9.9530	-1.6409
23	$C_{multiring}$	0.2306	****	****	****	0.4918	****	****	0.1846	-0.2729	-15.2261	-0.1495
24	aC-CH_m-aC (different rings) (m in 0..2)	0.0575	0.6827	0.0050	-18.7698	-0.2386	-0.0193	3.4750	-0.0300	-0.1612	9.3713	****
25	aC-$(CH_m$=$CH_n)$-aC (different rings) (m,n in 0..2)	0.2809	1.8492	0.0046	10.5243	0.7931	-0.9143	-5.3249	8.7846	0.5223	****	****
26	$(CH_m$=$C)_{cyc}$-CH=CH-$(C$=$CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	****	****	****
27	$(CH_m$=$C)_{cyc}$-CH_p-$(C$=$CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	0.8029	****	****
28	aC-CO-aC (different rings)	0.0515	0.5272	****	-29.1662	-0.3751	0.3143	18.2122	-0.0646	0.2438	****	****
29	aC-CH_m-CO-aC (different rings) (m in 0..2)	0.0515	****	****	****	-0.1937	****	****	****	0.2335	****	****
30	aC-CO-$(C$=$CH_n)_{cyc}$ (different rings) (n in 0..1)	-0.1303	****	****	****	-0.0075	****	****	****	0.2717	****	****
31	aC-CO-CO-aC (different rings)	****	****	****	****	2.3411	****	****	0.0829	0.9373	****	****
32	aC-CO_{cyc} (fused rings)	-0.1818	****	****	****	0.2248	****	46.3701	3.9705	0.1484	-3.1904	****
33	aC-CO-$(CH_n)_m$-CO-aC (different rings) (m>0; n in 0..2)	****	****	****	****	-1.1133	****	6.7784	****	****	****	****
34	aC-CO-CH_{ncyc} (different rings) (n in 0..1)	0.3352	****	****	****	-3.4715	****	****	****	0.7547	****	****

35	aC-CO-NH_n-aC (different rings) (*n* in 0..1)	****	****	****	****	-0.0041	****	****	0.0000	0.5650	****	****
36	aC-NH_nCONH_m-aC (different rings) (*n,m* in 0..1)	-0.0126	****	****	****	4.9128	****	****	0.0000	0.2396	****	****
37	aC-CO-N_{cyc} (different rings)	****	****	****	****	-0.0718	****	****	****	0.1221	****	****
38	aC-S_{cyc} (fused rings)	-0.4540	12.1354	0.0192	-15.5542	-0.3194	27.3990	29.6519	-0.0237	0.1765	****	****
39	aC-S-aC (different rings)	0.2799	****	****	****	0.3704	****	****	****	0.4717	****	****
40	aC-PO_n-aC (different rings) (*n* in 0..4)	****	****	****	****	****	****	****	****	****	****	****
41	aC-SO_n-aC (different rings) (*n* in 1..4)	****	****	****	****	1.7233	****	****	0.0046	0.0554	****	****
42	aC-NH_{ncyc} (fused rings) (*n* in 0..1)	-0.1551	3.5500	0.0282	****	0.2224	1.2664	2.2343	0.3632	0.2478	****	****
43	aC-NH-aC (different rings)	0.1996	5.7738	0.0187	-14.5590	0.1797	0.0945	-20.0410	0.2766	0.4814	4.7547	****
44	aC-$(C{=}N)_{cyc}$ (different rings)	-0.2846	****	****	****	-2.3712	****	****	****	-0.1555	****	****
45	aC-$(N{=}CH_n)_{cyc}$ (fused rings) (*n* in 0..1)	-0.4169	****	****	****	1.0888	****	-27.9439	-0.2137	0.0826	****	****
46	aC-$(CH_n{=}N)_{cyc}$ (fused rings) (*n* in 0..1)	****	****	****	****	0.5835	****	43.8690	****	-0.5766	****	****
47	aC-O-CH_n-aC (different rings) (*n* in 0..2)	****	****	****	****	-0.0219	****	****	****	-0.1850	****	****
48	aC-O-aC (different rings)	-0.2287	-3.7437	0.0113	-19.0021	-1.4721	0.6950	-8.8830	-6.7204	0.3203	0.0000	****
49	aC-CH_n-O-CH_m-aC (different rings) (*n,m* in 0..2)	-0.0478	-0.1595	0.0027	17.6609	-1.1688	0.0026	-	-0.0184	-0.5349	****	****
50	aC-O_{cyc} (fused rings)	-0.2868	5.1557	****	-12.7456	-0.1806	****	18.8585	0.0434	0.4148	****	****
51	AROM.FUSED[2]	-0.1599	-3.9457	0.0003	32.3905	-0.7199	0.6734	22.4274	0.0423	0.1818	7.8222	1.6962
52	AROM.FUSED[2]s^1	-0.1605	2.1630	-0.0026	-1.0416	-0.4454	0.0301	5.2245	-0.0381	0.2262	-7.5415	-1.9103
53	AROM.FUSED[2]s^2	-0.1829	-1.1898	-0.0012	9.7587	-0.2170	0.0805	12.3768	0.0442	0.1891	-3.5529	****
54	AROM.FUSED[2]s^2s^3	-0.0200	25.0034	****	****	0.6138	****	****	2.4113	0.1556	****	****
55	AROM.FUSED[2]s^1s^4	-0.2583	****	****	****	-0.2159	****	****	-0.4767	0.3293	****	5.2494
56	AROM.FUSED[2]s^1s^2	-0.0595	****	****	****	-0.8992	****	****	****	0.1684	****	****
57	AROM.FUSED[2]s^1s^3	-0.2059	26.5956	****	****	1.4041	****	****	****	0.3879	****	5.0594
58	AROM.FUSED[3]	-0.1825	-2.9209	0.0043	79.0621	-0.4008	0.3666	20.4876	0.1279	0.3070	****	****
59	AROM.FUSED[4a]	-0.0601	20.4501	0.0114	3.3996	-0.8482	8.4265	35.3367	9.2917	0.3796	****	****
60	AROM.FUSED[4a]s^1	0.1517	****	****	****	0.3308	****	41.4570	****	0.2155	****	****
61	AROM.FUSED[4a]s^1s^4	****	****	****	****	-2.0352	****	45.3846	****	1.1572	****	****
62	AROM.FUSED[4p]	0.0119	17.0823	0.0057	3.6510	0.0641	-23.6409	-7.9398	-4.0340	0.4212	****	-2.5282
63	AROM.FUSED[4p]s^3s^4	****	****	****	****	-0.5200	****	****	****	****	****	****
64	PYRIDINE.FUSED[2]	-0.1506	2.4430	-0.0027	15.9070	-0.6447	2.9399	12.2098	-5.8446	-0.2425	-1.8849	2.2059
65	PYRIDINE.FUSED[2-iso]	-0.2055	15.2077	-0.0067	6.1270	-0.0190	0.1719	12.0758	-4.4053	-0.1382	****	-3.9641
66	PYRIDINE.FUSED[4]	0.1771	****	0.0016	13.9203	-0.3820	-0.1599	34.6345	-0.1184	-0.0428	****	****
67	aC-N-CH_{cyc} (different rings)	****	****	****	****	1.7662	****	****	****	-0.3587	****	****
68	N multiring	-0.4064	****	****	****	0.7282	****	12.9731	****	-0.5770	****	****
69	N_{cyc}-$(CH_2)_3$-N_{cyc} (different rings)	****	****	****	****	-2.1419	****	****	****	-0.1105	****	****
70	aC-$COCH_2CH_2$-aC (different rings)	****	****	****	****	****	****	****	****	-0.6052	****	****
71	aC-O-$(CH_2)_2$-N_{cyc} (different rings)	****	****	****	****	****	****	****	****	-0.4688	****	****
72	aC-CH(OH)$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	-1.1757	****	****

73	Ncyc-$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	0.3889	****	****
74	aC-$CONHCH_2$-CH_{cyc} (different rings)	****	****	****	****	-1.2008	****	****	****	-0.3023	****	****

[a] The symbols T_{b3k}, T_{c3k}, P_{c3k}, V_{c3k}, T_{m3k}, $\Delta_f G^o_{3k}$, $\Delta_f H^o_{gas3k}$, $\Delta_{fus}H_{3k}$, $LogK_{ow3k}$, F_{p3k}, and $\Delta_{vap}H^o_{3k}$, represent the contributions (E_k) of the third-order groups for the corresponding properties

Table A7 (Continued). MG method based property models analysed using simultaneous regression method: Third-order groups and their contributions [a] along with sample assignments for the properties÷ $\Delta_{vap}H$, $\Delta_{vap}S$, δ_D, δ_P, δ_H, δ, AiT, ω and V_m

	Group	$\Delta_{vap}H_{3k}$	$\Delta_{vap}S_{3k}$	δ_{D3k}	δ_{P3k}	δ_{H3k}	δ_{3k}	$AiT1_{3k}$		ω_{3k}	V_{m3k}
1	HOOC-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	0.2747	3.3855	0.2768	-1.1559	-0.5205	-161.7973	0.0002	-0.0100
2	NH_n-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	****	****
3	NH_2-$(CH_n)_m$-OH (m>2, n in 0..2)	****	****	****	****	****	5.5667	****	****	0.0052	-0.0030
4	OH-$(CH_n)_m$-OH (m>2, n in 0..2)	0.2692	-10.1635	0.2783	2.3450	-2.4542	5.1190	-0.2542	-74.5818	0.0025	-0.0019
5	OH-$(CH_p)_k$-O-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****	****
6	OH-$(CH_p)_k$-S-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****	****
7	OH-$(CH_p)_k$-NH_x-$(CH_n)_m$-OH (m,k>0; p,n,x in 0..2)	****	****	****	****	****	****	****	****	****	****
8	CH_p-O-$(CH_n)_m$-OH (m>2; n,p in 0..2)	****	****	****	****	****	****	****	****	****	****
9	NH_2-$(CH_n)_m$-NH_2 (m>2; n in 0..2))	4.0694	2.9020	****	****	****	-0.4621	****	****	-0.0012	0.0211
10	NH_k-$(CH_n)_m$-NH_2 (m>2; k in 0..1; n in 0..2)	****	****	****	****	****	****	****	****	****	****
11	SH-$(CH_n)_m$-SH (m>2; n in 0..2)	****	****	****	****	****	****	****	****	****	****
12	NC-$(CH_n)_m$-CN (m>2)	****	****	****	****	****	-1.8285	2.0723	0.9993	0.0001	0.0004
13	COO-$(CH_n)_m$-OOC (m>2; n in 0..2)	****	****	****	****	****	****	****	****	****	****
14	aC-$(CH_n$=$CH_m)_{cyc}$ (fused rings) (n,m in 0..1)	****	****	0.9160	0.4886	0.9507	0.6218	****	****	0.0005	-0.0017
15	aC-aC (different rings)	****	****	2.7334	0.2887	-1.0851	1.4608	4.0493	-320.2954	0.0007	0.0070
16	aC-CH_{ncyc} (different rings) (n in 0..1)	****	****	0.2818	0.8041	-3.5861	1.5947	****	****	0.0092	0.0017
17	aC-CH_{ncyc} (fused rings) (n in 0..1)	0.9684	2.3237	3.2495	-0.8180	0.9306	1.4548	0.2251	47.7115	0.0003	-0.0025
18	aC-$(CH_n)_m$-aC (different rings) (m>1; n in 0..2)	****	****	****	****	****	1.3653	0.0360	-69.3483	-0.0027	-0.0037
19	aC-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****	****
20	CH_{cyc}-CH_{cyc} (different rings)	****	****	-0.7541	5.1632	-9.7602	****	0.2404	-27.0233	0.0166	****
21	CH_{cyc}-$(CH_n)_m$-CH_{cyc}(different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****	****
22	$CH_{multiring}$	0.0507	0.1656	4.4252	2.5938	-0.4329	0.0541	-0.0483	-18.9012	0.0080	-0.0002
23	$C_{multiring}$	0.0713	0.1124	6.6721	3.1420	-0.6190	0.3988	-0.1835	-51.2394	-0.0029	-0.1393
24	aC-CH_m-aC (different rings) (m in 0..2)	****	****	-0.4437	0.0714	-2.3539	0.7082	2.2165	-36.4427	0.0000	0.0001
25	aC-$(CH_m$=$CH_n)$-aC (different rings) (m,n in 0..2)	****	****	****	****	****	0.1270	****	****	-0.0002	-0.0043
26	$(CH_m$=$C)_{cyc}$-CH=CH-$(C$=$CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	****	****
27	$(CH_m$=$C)_{cyc}$-CH_p-$(C$=$CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	****	****
28	aC-CO-aC (different rings)	****	****	****	3.3685	3.1782	****	****	****	-0.0023	-0.0073
29	aC-CH_m-CO-aC (different rings) (m in 0..2)	****	****	-0.7702	0.0619	-1.7604	-0.7056	****	****	****	****
30	aC-CO-$(C$=$CH_n)_{cyc}$ (different rings) (n in 0..1)	****	****	****	****	****	****	****	****	****	****
31	aC-CO-CO-aC (different rings)	****	****	****	****	****	****	****	****	****	****

32	aC-CO_{cyc} (fused rings)	****	****	1.5650	-1.4124	-0.2357	-1.7052	-0.1089	-1.8369	-0.0008	0.0043
33	aC-CO-$(CH_n)_m$-CO-aC (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****	****
34	aC-CO-CH_{ncyc} (different rings) (n in 0..1)	****	****	****	****	****	****	****	****	****	****
35	aC-CO-NH_n-aC (different rings) (n in 0..1)	****	****	****	****	****	****	****	****	****	****
36	aC-NH_nCONH_m-aC (different rings) (n,m in 0..1)	****	****	****	****	****	****	****	****	****	****
37	aC-CO-N_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****
38	aC-S_{cyc} (fused rings)	****	****	2.5981	-9.2354	-1.4420	-0.0268	****	****	0.0002	0.0033
39	aC-S-aC (different rings)	****	****	****	****	****	****	****	****	****	****
40	aC-PO_n-aC (different rings) (n in 0..4)	****	****	****	****	****	****	****	****	****	****
41	aC-SO_n-aC (different rings) (n in 1..4)	****	****	-0.6929	-2.5668	-2.6053	****	****	****	****	****
42	aC-NH_{ncyc} (fused rings) (n in 0..1)	****	****	3.1324	0.9256	0.7173	0.2049	****	****	0.0007	-0.0013
43	aC-NH-aC (different rings)	****	****	-0.4927	-0.1207	-3.5536	1.0732	2.2137	59.2008	-0.0001	0.0039
44	aC-$(C=N)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	****	****
45	aC-$(N=CH_n)_{cyc}$ (fused rings) (n in 0..1)	****	****	1.2840	-0.7559	0.2899	****	****	****	-0.0009	0.0264
46	aC-$(CH_n=N)_{cyc}$ (fused rings) (n in 0..1)	****	****	3.1477	-3.5059	-0.0743	****	****	****	****	****
47	aC-O-CH_n-aC (different rings) (n in 0..2)	****	****	****	****	****	****	****	****	****	****
48	aC-O-aC (different rings)	****	****	-0.9055	-2.1463	-1.3514	-2.2322	0.8996	-28.9901	-0.0033	0.0017
49	aC-CH_n-O-CH_m-aC (different rings) (n,m in 0..2)	****	****	0.1723	5.3792	1.9259	-1.0428	****	****	0.0039	0.0010
50	aC-O_{cyc} (fused rings)	****	****	2.1052	-0.3721	-0.5115	2.6106	****	****	-0.0034	0.0042
51	AROM.FUSED[2]	-1.6909	-4.3390	0.1979	-1.5460	2.6294	0.2056	34.9646	74.0008	0.0033	0.0078
52	AROM.FUSED[2]s^1	5.0866	12.5449	0.7197	0.5326	0.8405	0.6312	34.9330	49.3633	0.0028	-0.0069
53	AROM.FUSED[2]s^2	****	****	-0.4481	-1.1340	0.0189	0.5135	34.8997	10.7121	0.0033	0.0063
54	AROM.FUSED[2]s^2s^3	****	****	****	****	****	****	****	****	****	****
55	AROM.FUSED[2]s^1s^4	****	****	****	****	****	****	****	****	****	****
56	AROM.FUSED[2]s^1s^2	****	****	****	****	****	****	****	****	****	****
57	AROM.FUSED[2]s^1s^3	****	****	****	****	****	****	****	****	****	****
58	AROM.FUSED[3]	****	****	****	****	****	-0.2831	****	****	0.0031	0.0094
59	AROM.FUSED[4a]	****	****	****	****	****	0.2376	69.6499	-16.3760	0.0058	0.0024
60	AROM.FUSED[4a]s^1	****	****	****	****	****	****	****	****	****	****
61	AROM.FUSED[4a]s^1s^4	****	****	****	****	****	****	****	****	****	****
62	AROM.FUSED[4p]	****	****	****	****	****	-0.5205	****	****	0.0044	-0.0143
63	AROM.FUSED[4p]s^3s^4	****	****	****	****	****	****	****	****	****	****
64	PYRIDINE.FUSED[2]	****	****	0.5017	-2.9280	-0.7278	1.1618	34.8229	4.3572	0.0024	-0.0034
65	PYRIDINE.FUSED[2-iso]	****	****	****	****	****	1.8826	****	****	0.0009	-0.0026
66	PYRIDINE.FUSED[4]	****	****	2.0792	-0.6601	-3.0822	1.9132	****	****	0.0047	0.0077

67	aC-N-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****
68	N multiring	****	****	4.0295	2.7546	3.2021	-1.2129	****	****	0.0024	-0.0001
69	N_{cyc}-$(CH_2)_3$-N_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****
70	aC-$COCH_2CH_2$-aC (different rings)	****	****	****	****	****	****	****	****	****	****
71	aC-O-$(CH_2)_2$-N_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****
72	aC-CH(OH)$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****
73	Ncyc-$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****
74	aC-$CONHCH_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****

[a] The symbols $\Delta_{vap}H_{3k}$, $\Delta_{vap}S_{3k}$, $\delta_{D3k,}$ $\delta_{P3k,}$ δ_{H3k}, Ait_{a3k}, Ait_{b3k}, ω_{3k}, and V_{m3k} represent the contributions (E_k) of the third-order groups for the corresponding properties.

Table A8. CI method based property models: atom contributions and model constants for the properties÷ T_b, T_c, P_c, V_c, T_m, $\Delta_f G^o$, $\Delta_f H^o_{gas}$, $\Delta_{fus}H$, $LogK_{ow}$, F_p, $\Delta_{vap}H^o$, $\Delta_{vap}H$, ω, and V_{mo}

	Properties												
Parameter	T_b	T_c	P_c	V_c	T_m	$\Delta_f G^o$	$\Delta_f H^o_{gas}$	$\Delta_{fus}H$	$LogK_{ow}$	F_p	$\Delta_{vap}H^o$	$\Delta_{vap}H$	ω	V_m
a(H)	-0.0556	-0.8787	0.0001	5.8893	-0.0979	-15.7725	-33.1687	-0.0948	-0.1169	-0.3051	-0.3219	0.0296	0.2864	0.0024
a(Cl)	1.0949	10.7974	-0.0050	40.3898	1.3595	-49.9743	-62.7872	4.2353	0.1600	38.4976	7.3166	5.1743	0.7986	0.0093
a(Br)	1.8894	16.7432	-0.0241	46.8027	2.6293	-32.1809	-30.7824	5.2518	-0.0315	67.9870	12.5516	8.2325	0.9678	0.0074
a(F)	0.0435	0.3456	0.0065	19.1937	0.2735	-222.6764	-235.3690	0.6465	0.0263	-222.5988	0.7152	0.5996	0.8209	0.0069
a(I)	2.6207	26.3668	-0.0230	62.8942	3.6072	9.2200	12.9869	6.0726	-0.7851	11.3109	16.5039	11.5155	0.9914	0.0083
a(N)	1.3049	13.7498	-0.0032	38.8734	2.5741	70.1603	73.6307	4.8104	-0.4709	43.5384	9.6875	5.7976	1.0653	0.0031
a(O)	0.8702	8.4451	-0.0052	17.0554	1.8184	-168.2051	-168.2467	4.0107	-0.5536	37.6959	7.4939	5.2902	1.6640	0.0028
a(P)	1.0008	48.7891	0.0010	-87.5564	0.6917	****	-232.5842	-13.0005	-0.3660	****	-243.5295	-243.5295	0.9977	0.0135
a(S)	1.5915	17.5489	-0.0117	29.5833	1.7393	-1.2219	5.8252	3.8558	0.0056	44.2641	9.0847	5.7623	0.9512	0.0073
a(C)	0.7699	7.1137	0.0020	33.2804	0.7636	35.4469	36.9694	2.3025	0.2477	18.9312	5.7963	3.1679	0.5187	0.0075
a(Si)	0.6873	10.4669	0.0232	0.0010	-0.4078	****	****	****	0.7026	****	****	****	0.9948	0.0284
b	-0.5468	-5.5389	0.0063	1.2217	-0.0151	-20.5110	-5.5558	-4.6180	0.2440	-27.0324	-6.8230	-4.4857	0.5900	0.0045
c	0.2342	1.9755	0.0032	10.3555	-0.2487	15.6338	10.1904	3.7133	0.0698	5.8538	3.8729	2.0165	0.7095	0.0011
d	1.5792	0.7117	0	-6.4593	0	61.7733	28.5860	0.9497	0.0850	-3.6112	36.3136	39.0552	0.7890	0.0041

Appendix B

Model performance statistics, group-contribution tables and atom-contribution tables for GC^+ models for the estimation of environment-related properties of pure components.

Table B1. Performance of MG Method Based Property Models Analysed by Simultaneous Regression Method

Table B2. MG Method Based Property Models Analysed Using Step-Wise Regression Method: First-Order Groups and their Contributions for the Environmental-Related Properties

Table B3. MG Method Based Property Models Analysed Using Step-Wise Regression Method: Second-Order Groups and their Contributions for the Environmental-Related Properties

Table B4. MG Method Based Property Models Analysed Using Step-Wise Regression Method: Third-Order Groups and their Contributions for the Environmental-Related Properties

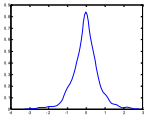
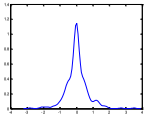
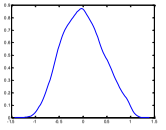
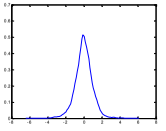
Table B5. MG Method Based Property Models Analysed Using Simultaneous Regression Method: First-Order Groups and their Contributions for the Environmental-Related Properties

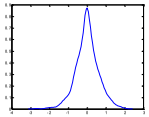
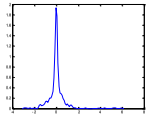
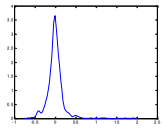
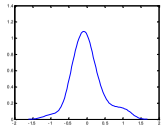
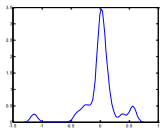
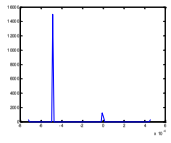
Table B6. MG Method Based Property Models Analysed Using Simultaneous Regression Method: Second-Order Groups and their Contributions for the Environmental-Related Properties

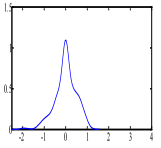
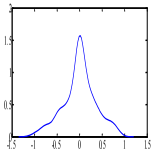
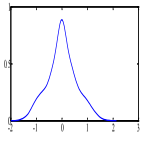
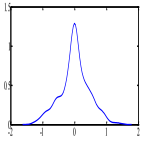
Table B7. MG Method Based Property Models Analysed Using Simultaneous Regression Method: Third-Order Groups and their Contributions for the Environmental-Related Properties

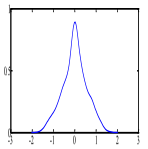
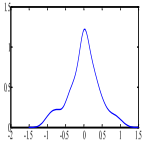
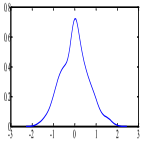
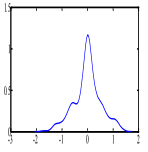
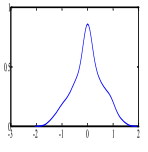
Table B8. CI Method Based Property Models: Atom Contributions and Model Constants for the Environmental-Related Properties

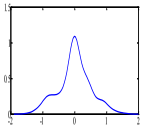
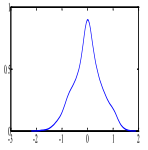
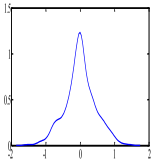
Table B1. Performance of MG Method Based Models for Environmental-Related Properties Analysed by Simultaneous Regression Method

sl. no.	property	L.H.S. of MG method based property prediction model	MG group-contribution model $f(X)=\sum_i N_i C_i+\sum_j M_j D_j+\sum_k E_k O_k$									
		$f(X)$	N	ν	R^2	residual distribution plot	P_{rc} (± 1%)	P_{rc} (± 5%)	P_{rc} (± 10%)	SD	AAE	ARE [a]
1	fathead minnow 96-hr LC_{50} (LC_{50} (FM)) in mol/lit	$-\text{LogLC}_{50}(\text{FM})+\text{FM}_0$	809	541	0.81		9.64	33.00	57.11	0.63	0.45	20.14
2	*daphnia magna* 48-hr LC_{50} (LC_{50} (DM)) in mol/lit	$-\text{LogLC}_{50}(\text{DM})+\text{DM}_0$	320	124	0.87		28.75	48.13	67.50	0.61	0.38	11.19
3	oral Rat LD_{50} (LD_{50}) in mol/kg	$-\text{LogLD50}-A_{LD50}-B_{LD50}\text{MW}$	5995	5617	0.74		1.87	7.96	14.18	0.42	0.35	16.09
4	aqueous solubility (LogW_S) in gm/lit	$\log(W_S)-A_{W_S}-B_{W_S}\text{MW}$	4681	4311	0.79		3.23	14.80	28.97	0.97	0.71	----

5	bioconcentration factor (BCF)	LogBCF	662	423	0.80		9.37	20.85	30.36	0.60	0.44	----
6	permissible Exposure Limit (PEL)) in mol/lit	-LogPEL	425	239	0.78		23.53	48.47	65.88	0.72	0.38	11.02
7	photochemical Oxidation Potential (PCO)	-LogPCO	639	488	0.86		5.95	18.31	28.64	0.20	0.12	6.60
8	global Warming Potential (GWP)	LogGWP	51	31	0.87		15.69	37.25	56.86	0.41	0.29	11.57
9	ozone Depletion Potential (ODP)	LogODP	28	12	0.89		17.86	21.43	28.5	0.30	0.16	----
10	acidification Potential (ODP)	LogAP	10	1	1.0		100.0	--	--	3.4E-04	2.1E-4	----

11	emission to urban air (EUA_C) in cases/kg emitted (carcinogenic)	$-\text{Log}(EUA_C)+A_{EUA_C}$	456	214	0.84		23.03	48.90	75.22		0.51	0.36	7.42
12	emission to urban air (EUA_{NC}) in cases/kg emitted (non-carcinogenic)	$-\text{Log}(EUA_{NC})+A_{EUA_{NC}}$	341	128	0.88		27.57	60.70	84.46		0.36	0.26	4.87
13	emission to continental rural air (ERA_C) in cases/kg emitted (carcinogenic)	$-\text{Log}(ERA_C)+A_{ERA_C}$	470	229	0.83		20.43	45.53	71.91		0.56	0.42	7.18
14	emission to continental rural air (ERA_{NC}) in cases/kg emitted (non-carcinogenic)	$-\text{Log}(ERA_{NC})+A_{ERA_{NC}}$	349	134	0.87		26.36	55.30	80.52		0.45	0.32	5.50

15	emission to continental fresh water (EFW_C) in cases/kg emitted (carcinogenic)	$-\text{Log}(EFW_C)+A_{EFW_C}$	472	230	0.83		20.34	41.95	68.43	0.55	0.41	8.87
16	emission to continental fresh water (EFW_{NC}) in cases/kg emitted (non-carcinogenic)	$-\text{Log}(EFW_{NC})+A_{EFW_{NC}}$	345	131	0.89		23.19	52.75	77.97	0.42	0.31	6.33
17	emission to continental sea water (ESW_C) in cases/kg emitted (carcinogenic)	$-\text{Log}(ESW_C)+A_{ESW_C}$	477	235	0.87		22.22	46.96	76.10	0.63	0.48	6.72
18	emission to continental sea water (ESW_{NC}) in cases/kg emitted (non-carcinogenic)	$-\text{Log}(ESW_{NC})+A_{ESW_{NC}}$	360	146	0.91		23.33	58.06	80.56	0.52	0.38	6.06
19	emission to continental natural soil (ENS_C) in cases/kg emitted (carcinogenic)	$-\text{Log}(ENS_C)+A_{ENS_c}+B_{ENS_c}\times MW$	460	219	0.84		24.13	50.00	76.52	0.52	0.38	6.22

20	emission to continental natural soil (ENS_{NC}) in cases/kg emitted (non-carcinogenic)	$-Log(ENS_{NC})+A_{ENS_{NC}}$	362	148	0.85		23.48	55.25	77.90	0.51	0.37	5.91
21	emission to continental agri-cultural soil (EAS_C) in cases/kg emitted (carcinogenic)	$-Log(EAS_C)+A_{EAS_C}$	470	228	0.84		22.34	47.23	73.83	0.53	0.40	6.91
22	emission to continental agri-cultural soil (EAS_{NC}) in cases/kg emitted (non-carcinogenic)	$-Log(EAS_{NC})+A_{EAS_{NC}}$	352	138	0.87		23.86	56.53	82.39	0.44	0.32	5.30

[a] ARE is not defined for LogWs, BCF, ODP and AP since these properties have both positive and negative values.

Table B2. MG Method Based Property Models Analysed Using Step-Wise Regression Method: First-Order Groups and their Contributions [a] for the Properties÷ LC_{50}(FM), LC_{50}(DM), LD_{50}, $LogW_s$, BCF, PEL(OSHA-TWA), PCO, GWP, ODP, and AP.

	Group	$LC50(FM)_{1i}$	$LC50(DM)_{1i}$	$LD50_{1i}$	$LogWs_{1i}$	BCF_{1i}	PEL_{1i}	PCO_{1i}	GWP_{1i}	ODP_{1i}	AP_{1i}
1	CH_3	0.0972	-0.0386	-0.0742	-5.2494	0.6657	0.7723	0.1227	0.3880	-0.9453	-0.1290
2	CH_2	0.2885	0.1710	0.0223	-5.0706	0.0948	0.0727	0.0463	-1.0699	****	****
3	CH	0.2441	-0.1654	0.1335	-4.8948	-0.3921	-0.6557	-0.0790	****	****	****
4	C	-0.3822	0.4640	0.2641	-4.6277	-0.9137	-1.3404	-0.0434	****	****	****
5	CH_2=CH	1.0340	0.1698	0.1087	-9.6240	0.7712	2.2638	-0.2572	****	****	****
6	CH=CH	0.3890	-0.2512	0.0977	-9.5597	****	0.7658	-0.5513	****	****	****
7	CH_2=C	0.5436	0.5657	0.1358	-9.4175	-0.1358	0.1652	-0.3932	****	****	****
8	CH=C	0.5902	0.6270	0.2376	-9.4095	-0.0639	-0.7468	-0.6306	****	****	****
9	C=C	0.8302	0.2129	0.3451	-9.0081	-0.1808	-1.2669	-0.6010	****	****	****
10	CH_2=C=CH	****	****	****	****	****	****	****	****	****	****
11	CH_2=C=C	****	****	****	****	****	****	****	****	****	****
12	C=C=C	****	****	****	****	****	****	****	****	****	****
13	CH≡C	0.7491	****	0.0609	-9.0073	****	3.0138	-0.3025	****	****	****
14	C≡C	1.6682	-0.2044	0.3876	-8.6808	****	****	-0.9074	****	****	****
15	aCH	0.1530	0.0743	0.0230	-4.6135	0.2561	0.4045	0.0137	****	****	****
16	aC fused with aromatic ring	0.1782	0.2851	0.0398	-4.7032	-0.0341	0.2664	0.0053	****	****	****
17	aC fused with non-aromatic ring	0.6141	0.4121	0.1363	-4.5533	0.1641	0.3003	-0.0341	****	****	****
18	aC except as above	0.3150	0.3864	0.0501	-4.5067	-0.0071	-0.0178	0.8000	****	****	****
19	aN in aromatic ring	-0.1501	-0.2125	0.0755	-4.4745	0.0130	1.4732	****	****	****	****
20	aC-CH_3	0.4050	0.4345	0.0699	-9.6580	0.4333	0.5522	-0.1132	****	****	****
21	aC-CH_2	0.3350	0.2087	0.0821	-9.3278	-0.1904	-0.3934	-0.2018	****	****	****
22	aC-CH	0.7893	0.9717	0.4545	-9.0233	-0.9151	-1.1787	-0.2998	****	****	****
23	aC-C	0.5639	1.0938	0.4264	-8.7328	-0.9509	-0.9892	-0.3527	****	****	****
24	aC-CH=CH_2	0.8110	0.0652	0.1599	-14.0398	0.8698	0.5330	0.3822	****	****	****
25	aC-CH=CH	0.5427	****	0.0900	-14.4421	0.1297	****	0.8379	****	****	****
26	aC-C=CH_2	****	****	-0.6814	-14.2994	0.3000	-0.4032	0.1949	****	****	****
27	aC-C≡CH	****	****	****	-13.6670	****	****	****	****	****	****
28	aC-C≡C	****	****	-0.3680	-14.4769	****	****	****	****	****	****
29	OH	-0.6115	-0.8815	-0.1955	-5.1862	-0.0340	1.3612	0.0359	****	****	****
30	aC-OH	0.2670	0.2670	0.0705	-9.6989	-0.2700	1.2393	0.3113	****	-1.5842	-0.0769
31	COOH	-0.1104	-0.3982	0.0320	-14.8398	-0.8830	2.3281	-0.0379	****	****	****
32	aC-COOH	0.0172	-1.6411	0.0391	-19.4187	-0.9232	1.1988	****	****	****	****
33	CH_3CO	0.0835	-1.4018	-0.0172	-14.2841	0.6654	1.4016	0.1409	****	****	****
34	CH_2CO	-0.5508	****	0.1931	-14.6443	-1.4842	1.2601	-0.0515	****	****	****
35	CHCO	****	****	0.4130	-15.1578	****	****	0.0505	****	****	****
36	CCO	****	****	1.1655	-14.0177	****	0.1081	****	****	****	****
37	aC-CO	0.2699	0.4353	0.2190	-14.1140	-0.9893	0.9671	****	****	****	****
38	CHO	0.6008	0.7531	-0.1338	-9.3615	-0.3560	2.3662	-0.1855	****	****	****
39	aC-CHO	0.8678	0.9419	-0.0626	-14.4954	-1.3452	****	****	****	****	****

40	CH_3COO	0.4393	-0.6598	-0.1734	-19.9009	0.3988	1.2544	0.3858	****	****	****
41	CH_2COO	0.1823	0.4007	-0.0357	-19.7361	0.2378	1.6798	0.3021	****	****	****
42	CHCOO	-0.0242	****	0.1329	-19.3313	****	****	0.1068	****	****	****
43	CCOO	-1.9110	****	0.3242	-19.3558	****	****	****	****	****	****
44	HCOO	****	****	-0.0612	-15.4178	****	1.5780	0.8282	****	****	****
45	aC-COO	0.4382	0.0355	-0.0431	-19.3244	-1.1633	0.5326	-0.0378	****	****	****
46	aC-OOCH	****	****	-0.2398	****	****	****	****	****	****	****
47	aC-OOC	0.9868	****	0.4845	-19.5174	-2.9904	0.9678	****	****	****	****
48	COO except as above	0.4138	1.3113	0.0657	-14.9256	-0.5196	0.4342	-0.2744	****	****	****
49	CH_3O	-0.5209	0.7930	-0.0259	-10.1615	0.2530	2.1251	0.1499	0.1245	****	****
50	CH_2O	-0.2160	0.2341	0.0974	-9.9546	-0.1060	0.9276	-0.1226	****	****	****
51	CH-O	-0.8189	****	0.4987	-9.8757	-0.3114	-0.7462	-0.2064	-1.8521	****	****
52	C-O	****	****	0.2275	-9.2509	-0.7149	****	****	****	****	****
53	aC-O	0.0896	0.7875	0.1839	-9.6929	-0.3264	0.4621	-0.4603	****	****	****
54	CH_2NH_2	-0.0246	0.1531	0.0450	-9.5161	0.2885	1.9265	-0.3290	****	****	****
55	$CHNH_2$	-0.0983	0.9878	0.3764	-10.1966	****	2.1480	-0.5376	****	****	****
56	CNH_2	0.5803	****	0.4151	-8.7982	****	****	****	****	****	****
57	CH_3NH	-0.4593	0.1983	-0.0593	-9.5708	****	2.8505	0.0046	****	****	****
58	CH_2NH	0.0524	0.0553	0.2571	-9.2845	-0.3334	1.2126	-0.0060	****	****	****
59	CHNH	****	-0.3115	0.2506	-8.8957	****	1.2708	****	****	****	****
60	CH_3N	0.1273	-0.9599	0.3338	-9.5015	-0.4163	1.1981	-0.4139	****	****	****
61	CH_2N	-0.3121	-1.4049	0.4337	-9.3647	-1.4607	0.2724	-0.3483	****	****	****
62	aC-NH_2	0.0721	1.0918	0.1189	-9.8417	-0.3313	2.0982	****	****	****	****
63	aC-NH	0.3195	0.4657	0.1796	-9.7080	-0.4387	0.7464	****	****	****	****
64	aC-N	0.0691	0.5173	0.2682	-9.5308	-0.6948	-0.7120	****	****	****	****
65	NH_2 except as above	0.1130	0.2107	0.0130	-5.2748	-0.1643	2.0449	-0.4319	****	****	****
66	CH=N	2.4719	0.0832	0.1706	-9.5889	0.1155	****	****	****	****	****
67	C=N	0.9533	1.1620	0.2576	-8.8434	-1.0204	****	****	****	****	****
68	CH_2CN	-0.0683	****	-0.0168	-13.5365	0.3979	2.0501	****	****	****	****
69	CHCN	2.0149	****	1.0312	-13.2307	-0.4915	1.9526	****	****	****	****
70	CCN	1.0777	****	0.7798	-12.9126	****	0.7938	****	****	****	****
71	aC-CN	0.0788	0.6363	0.1041	-13.5595	-0.3186	****	****	****	****	****
72	CN except as above	0.5895	0.2307	0.3921	-9.2357	1.2018	2.5607	0.4704	****	****	****
73	CH_2NCO	****	****	0.0453	****	****	3.1956	****	****	****	****
74	CHNCO	****	****	****	****	****	****	****	****	****	****
75	CNCO	****	****	****	****	****	****	****	****	****	****
76	aC-NCO	****	****	-0.1489	****	****	2.0353	-0.0504	****	****	****
77	$CH2NO_2$	****	****	0.1374	-20.3294	****	1.8813	1.0797	****	****	****
78	$CHNO_2$	****	****	0.6535	-20.6143	-1.0104	1.4511	****	****	****	****
79	CNO_2	1.4039	****	1.1115	-18.8805	****	0.1482	****	****	****	****
80	aC-NO_2	0.7014	0.2756	0.2627	-20.3307	-0.0368	1.5088	1.3371	****	****	****
81	NO_2 except as above	1.0513	-1.3024	-0.0320	-16.4594	****	1.4137	****	****	****	-0.0775
82	ONO	****	****	0.8495	****	****	****	****	****	****	****
83	ONO_2	-0.0796	-0.3048	0.2883	-21.5057	****	2.0086	****	****	****	-0.0775
84	$HCON(CH_2)_2$	-0.2502	****	****	-22.5375	****	****	****	****	****	****
85	$HCONHCH_2$	****	****	****	-19.5759	****	****	****	****	****	****

86	$CONH_2$	-0.1815	1.2951	0.1678	-14.6677	****	3.1108	****	****	****	****
87	$CONHCH_3$	1.2999	-0.2366	1.0423	-19.5825	-0.6677	****	****	****	****	****
88	$CONHCH_2$	0.9036	3.7503	0.1446	-19.3632	-1.7679	****	****	****	****	****
89	$CON(CH_3)_2$	-0.9938	****	0.7419	-23.9555	1.7095	2.6238	****	****	****	****
90	$CONCH_3CH_2$	****	****	****	-23.7463	****	****	****	****	****	****
91	$CON(CH_2)_2$	-0.5947	-0.5260	0.3286	-23.5108	-0.6649	****	****	****	****	****
92	CONHCO	****	****	0.5830	-24.4140	****	****	****	****	****	****
93	CONCO	1.3874	****	0.4392	-23.7170	****	****	****	****	****	****
94	$aC\text{-}CONH_2$	-0.3693	-1.3986	0.2374	-19.1965	****	****	****	****	****	****
95	aC-NH(CO)H	-0.0863	****	0.2940	-18.3534	****	****	****	****	****	****
96	aC-N(CO)H	-0.3164	****	0.4427	****	****	****	****	****	****	****
97	aC-CONH	1.2356	-2.8641	0.0444	-18.6958	-1.8423	****	****	****	****	****
98	aC-NHCO	-0.1143	0.0285	0.0436	-18.9709	-1.2286	****	****	****	****	****
99	aC-(N)CO	0.1407	-0.0767	0.2094	-18.1536	-2.0069	****	****	****	****	****
100	NHCONH	****	****	0.1069	-20.4704	-3.0761	****	****	****	****	****
101	NH_2CONH	****	****	-0.0066	-20.2551	****	****	****	****	****	****
102	NH_2CON	****	****	0.7208	-19.6286	****	****	****	****	****	****
103	NHCON	-0.6989	****	0.5454	-19.0554	****	****	****	****	****	****
104	NCON	****	-1.2393	0.4743	-17.1827	****	****	****	****	****	****
105	$aC\text{-}NHCONH_2$	****	****	0.0291	-24.4105	****	****	****	****	****	****
106	aC-NHCONH	-0.5009	2.1330	0.0802	-24.9922	-1.4420	****	****	****	****	****
107	NHCO except as above	0.6549	-0.9079	0.1199	-14.2024	0.6908	****	****	****	****	****
108	CH_2Cl	0.6181	-0.1689	0.3896	-17.3359	0.4043	2.3309	0.3730	-0.2757	-1.4515	****
109	CHCl	0.1069	0.5411	0.2054	-17.1019	0.3504	-0.9263	0.6117	****	-1.3742	****
110	CCl	****	****	0.2454	-17.0334	****	****	****	****	****	****
111	$CHCl_2$	0.5662	0.3245	0.2650	-29.1152	0.4145	1.7348	1.5265	-0.0233	-1.7873	****
112	CCl_2	0.7493	****	-0.0580	-28.3154	0.8521	****	****	****	****	****
113	CCl_3	1.3517	1.2019	0.3039	-41.6097	0.9977	1.8541	2.7667	1.8173	-0.0399	-0.0132
114	CH_2F	****	****	2.1620	-10.7960	****	****	1.9489	0.8584	****	****
115	CHF	****	****	****	-9.2501	****	****	****	-0.6900	****	****
116	CF	****	****	****	****	****	****	****	****	****	****
117	CHF_2	1.0749	****	0.3702	-17.5991	****	****	1.4794	1.6870	-0.5628	****
118	CF_2	0.1261	****	0.1139	-18.6216	0.6323	-0.6817	****	-0.0026	-0.2492	****
119	CF_3	0.6351	-0.4425	1.3466	-24.1014	0.3278	1.3676	0.5828	2.1289	0.0883	****
120	CCl_2F	****	****	-0.2211	-36.0704	0.8196	0.5580	2.4794	2.1575	0.0131	0.1546
121	HCClF	****	****	-0.4169	-23.5774	****	0.0865	2.0994	0.9439	-1.4911	****
122	$CClF_2$	****	****	0.1143	-30.0684	0.8033	0.4950	2.4794	2.0786	-0.0409	****
123	aC-Cl	0.6323	0.4909	0.1991	-17.1355	0.6407	0.9796	0.6083	****	-0.2832	-0.0290
124	aC-F	0.3392	0.6346	0.2812	-11.0066	0.5743	****	0.3160	****	****	****
125	aC-I	0.9439	****	0.2841	-47.8855	0.7695	****	****	****	****	****
126	aC-Br	0.7252	0.8113	0.4920	-32.2539	0.3579	****	****	****	****	****
127	-I except as above	0.9066	-0.2044	0.4522	-43.7256	****	2.9327	****	-2.1289	****	****
128	-Br except as above	0.6769	0.4380	0.3970	-27.6030	0.5466	1.5345	0.9433	1.0645	0.4320	****
129	-F except as above	0.0814	0.1183	-0.0353	-6.6241	0.5813	1.2998	0.4224	1.9617	-0.0074	****
130	-Cl except as above	0.3776	0.1975	0.0133	-12.4658	0.5005	1.3093	0.5640	1.2757	-0.0056	-0.0682
131	CHNOH	0.7400	****	0.2486	-15.5688	****	****	****	****	****	****

132	CNOH	-0.4679	****	0.1320	-14.8149	-1.2471	****	****	****	****	****
133	aC-CHNOH	1.5961	****	0.9452	****	****	****	****	****	****	****
134	OCH_2CH_2OH	-0.7009	-1.2305	-0.1547	-19.6115	-0.2796	1.7163	0.0024	****	****	****
135	$OCHCH_2OH$	****	****	0.0366	-19.2869	****	-0.6939	-0.1438	****	****	****
136	OCH_2CHOH	0.1611	****	0.0275	-19.5555	****	0.2807	-0.1101	****	****	****
137	-O-OH	****	****	0.1372	-10.3102	****	****	0.6889	****	****	****
138	CH_2SH	1.5431	1.8996	0.4204	-16.6831	-2.2197	3.2668	****	****	****	****
139	CHSH	0.2299	****	0.0476	-16.7708	****	****	****	****	****	****
140	CSH	****	****	-0.1342	-14.8742	****	****	****	****	****	****
141	aC-SH	****	****	0.8198	-16.6828	****	3.3208	****	****	****	****
142	-SH except as above	0.7657	****	-0.0389	-10.6945	****	3.0292	****	****	****	****
143	CH_3S	0.4408	0.4059	0.4553	-16.4597	0.5000	2.6231	****	****	****	****
144	CH_2S	0.6252	1.0444	0.3767	-16.2943	-0.4380	****	****	****	****	****
145	CHS	0.6236	-0.3983	0.5770	-15.8430	-1.3246	-1.5055	****	****	****	****
146	CS	1.4408	****	0.2262	-16.8443	****	****	****	****	****	****
147	aC-S-	0.8459	0.3129	0.3336	-15.7817	-1.0938	****	****	****	****	****
148	SO	-1.3615	-2.3995	0.1737	-14.8690	-1.4619	1.7578	-0.5158	****	****	****
149	SO2	2.5611	****	0.3939	-21.8506	-0.5763	2.2691	****	****	****	****
150	SO_3 (sulfite)	****	****	-0.2231	-31.4604	****	****	****	****	****	****
151	SO_3 (Sulfonate)	0.4219	****	0.3187	-26.8991	****	****	****	****	****	****
152	SO_4 (Sulfate)	****	****	0.7967	-32.6492	****	2.8573	****	****	****	****
153	aC-SO	-0.5944	****	0.6778	-21.4786	-2.4005	****	****	****	****	****
154	aC-SO_2	0.2814	-1.0660	0.0021	-25.8954	-0.3242	****	****	****	****	****
155	PH (phosphine)	****	****	****	****	****	****	****	****	****	****
156	P (Phospine)	****	****	0.5950	****	****	****	****	****	****	****
157	PO_3 (Phospite)	****	****	0.4576	****	-2.1138	****	****	****	****	****
158	PHO_3 (Phosponate)	0.3208	****	0.0665	-25.9962	****	****	****	****	****	****
159	PO_3 (Phosponate)	-0.0390	3.2035	1.4048	-26.0106	-1.5601	****	****	****	****	****
160	PHO_4 (Phospate)	1.4767	****	0.0500	-30.9428	-2.2271	****	****	****	****	****
161	PO_4 (Phospate)	0.0857	1.6308	1.8110	-30.7800	-2.2559	****	****	****	****	****
162	aC-PO_4	1.9467	2.5522	1.1535	-36.5009	-1.3727	-0.4513	****	****	****	****
163	aC-P	****	****	0.6937	****	****	****	****	****	****	****
164	CO_3 (Carbonate)	****	****	0.4026	-20.0556	****	****	1.5672	****	****	****
165	C_2H_3O	0.8514	****	0.1473	-14.0019	****	1.0419	0.7183	****	****	****
166	C_2H_2O	-0.4972	****	-0.0635	-13.5542	****	****	****	****	****	****
167	C_2HO	****	****	-0.0608	-13.9569	****	****	****	****	****	****
168	CH_2 (cyclic)	0.0897	-0.0164	0.0305	-4.9128	0.2381	0.2678	0.0522	****	****	****
169	CH (cyclic)	0.3440	0.2230	0.1009	-4.7455	0.0526	-0.1033	0.0076	****	****	****
170	C (cyclic)	-0.0751	0.1635	0.2675	-4.2661	-0.4317	-0.6719	0.0405	****	****	****
171	CH=CH (cyclic)	0.4053	0.1881	0.0342	-9.1455	0.8670	1.1926	-0.1694	****	****	****
172	CH=C (cyclic)	0.7101	0.6201	0.1961	-9.0741	-0.9429	-0.3924	-0.5956	****	****	****
173	C=C (cyclic)	0.8299	0.9138	0.3052	-9.1726	-0.1101	-0.7986	****	****	****	****
174	CH_2=C (cyclic)	1.3284	0.3457	0.6753	-8.3601	1.2508	1.6196	-0.4417	****	****	****
175	NH (cyclic)	-0.5359	0.0690	0.1039	-5.0559	-0.4237	3.5925	****	****	****	****
176	N (cyclic)	-0.5352	0.4133	0.1509	-4.1475	-0.3082	0.2011	-0.1283	****	****	****
177	CH=N (cyclic)	-0.1997	0.0016	0.0247	-9.4054	0.0169	****	****	****	****	****

178	C=N (cyclic)	-0.0297	1.6297	0.1815	-9.3322	-0.6622	****	****	****	****	****
179	O (cyclic)	-0.0178	-0.3136	0.0485	-5.2684	-0.4086	1.0976	0.1839	****	****	****
180	CO (cyclic)	-0.0751	-0.0348	0.0362	-9.7030	-0.2482	1.4906	0.1464	****	****	****
181	S (cyclic)	0.5408	-0.0221	0.0662	-11.1395	0.2438	-3.4289	****	****	****	****
182	SO_2 (cyclic)	****	-0.1274	0.2491	-21.4443	-0.5767	****	****	****	****	****
183	>NH	0.0122	-0.2127	0.2383	-4.6764	-0.6793	****	****	****	****	****
184	-O-	-1.0802	1.9031	0.3655	-5.2880	-0.2204	-6.9696	****	****	****	****
185	-S-	0.3532	2.2864	0.0645	-11.7688	-0.1536	****	****	****	****	****
186	>CO	1.1595	0.4557	0.0463	-9.8009	0.4628	1.8939	0.1130	****	****	****
187	PO_2	****	****	0.1374	****	****	****	****	****	****	****
188	CH-N	****	1.7603	-0.1090	-8.9657	****	****	****	****	****	****
189	SiHO	****	****	****	****	****	****	****	****	****	****
190	SiO	****	****	-0.0007	-16.0461	-0.3101	-3.1492	****	****	****	****
191	SiH_2	****	****	-0.2393	****	****	****	****	****	****	****
192	SiH	****	****	****	****	****	****	****	****	****	****
193	Si	****	****	0.3458	-11.2041	-1.3606	****	****	****	****	****
194	$(CH_3)_3N$	****	****	****	****	****	****	****	****	****	****
195	N=N	0.4933	****	0.0539	-9.3563	****	****	****	****	****	****
196	C_{cyclic}=N-	0.8793	****	0.6162	-9.3709	-0.7675	****	****	****	****	****
197	C_{cyclic}=CH-	2.3085	2.1495	0.6016	-8.8475	-0.6405	1.3880	****	****	****	****
198	C_{cyclic}=NH	****	****	-0.0748	-9.8549	****	****	****	****	****	****
199	N=O	-0.3575	-0.1965	0.3113	-10.0720	-0.3770	****	****	****	****	-0.0775
200	C_{cyclic}=C	0.6756	****	0.5676	-8.5743	-0.9050	****	-0.1664	****	****	****
201	P=O	1.2123	2.2490	0.8100	-16.0285	-1.1477	****	****	****	****	****
202	N=N	****	-1.2397	0.3611	-10.7648	-1.4104	****	****	****	****	****
203	C=NH	-0.4605	****	-0.1942	-9.9767	-0.8936	****	****	****	****	****
204	>C=S	0.9273	0.3273	0.3458	-15.7902	0.0321	****	****	****	****	****
205	aC-CON	-0.1061	2.2769	0.2972	-17.6529	-2.6869	****	****	****	****	****
206	aC=O	-0.3836	****	-0.0132	-10.1596	0.1362	****	****	****	****	****
207	aN-	-0.6546	0.3255	0.1369	-4.1070	-1.0346	****	****	****	****	****
208	-Na	****	****	****	****	****	****	****	****	****	****
209	-K	****	****	****	****	****	****	****	****	****	****
210	HCONH	****	****	-0.1554	-14.4278	-0.7736	****	****	****	****	****
211	CHOCH	1.2130	0.4709	0.3383	-13.6062	****	****	****	****	****	****
212	C_2O	****	****	****	****	****	****	****	****	****	****
213	SiH_3	****	****	****	****	****	****	****	****	****	****
214	SiH_2O	****	****	****	****	****	****	****	****	****	****
215	CH=C=CH	****	****	****	-13.3744	****	****	****	****	****	****
216	CH=C=C	****	****	****	****	****	****	****	****	****	****
217	OP(=S)O	0.8141	2.9977	1.4164	-33.4285	-0.2438	0.5000	****	****	****	****
218	R	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
219	$CF_{2cyclic}$	****	****	0.2788	-17.8927	****	****	****	1.0010	****	****
220	CF_{cyclic}	****	****	0.3840	-11.5240	****	****	****	****	****	****

[a] The symbols $LC50(FM)_{1i}$, $LC50(DM)_{1i}$, $LD50_{1i}$, $LogWs_{1i}$, BCF_{1i}, PEL_{1i}, PCO_{1i}, GWP_{1i}, ODP_{1i},and AP_{1i} represent the contributions (C_i) of the first-order groups for the corresponding properties.

Table B2 (continued). MG Method Based Property Models Analysed Using Step-Wise Regression Method: First-Order Groups and their Contributions [a] for the Properties÷ EUA_C, EUA_{NC}, ERA_C, ERA_{NC}, EFW_C, EFW_{NC}, ESW_C, ESW_{NC}, ENS_C, ENS_{NC}, EAS_C, and EAS_{NC}.

	Group	$EUA_{C\,1i}$	$EUA_{NC\,1i}$	$ERA_{C\,1i}$	$ERA_{NC\,1i}$	$EFW_{C\,1i}$	$EFW_{NC\,1i}$	$ESW_{C\,1i}$	$ESW_{NC\,1i}$	$ENS_{C\,1i}$	$ENS_{NC\,1i}$	$EAS_{C\,1i}$	$EAS_{NC\,1i}$
1	CH_3	0.2812	0.0044	0.3053	0.1229	0.2626	0.1627	0.1362	-0.1075	0.2433	0.0393	0.1771	0.0598
2	CH_2	-0.0243	-0.0768	-0.0173	-0.0888	-0.0503	-0.1054	-0.1035	-0.1460	0.1514	0.0997	0.0888	0.0041
3	CH	-0.2311	0.0074	-0.0458	-0.0536	0.1164	-0.1383	0.1699	-0.1374	0.2068	0.2690	0.0505	0.0314
4	C	-0.2714	-0.2538	-0.8209	-0.8273	-0.5915	-0.9984	-0.6146	-0.5315	-0.6888	-0.1677	-0.7997	-0.5965
5	CH_2=CH	0.4720	-1.1792	0.7120	-0.7414	0.3625	-1.0817	-0.1128	-0.9796	0.4585	-0.4659	0.3632	-0.4702
6	CH=CH	0.0099	-0.6462	-0.0053	-0.3547	-0.5021	-0.3094	-0.6042	-0.1727	-0.2822	-0.3880	-0.3203	-0.3861
7	CH_2=C	-0.0002	-0.3848	0.9076	-0.0390	0.4974	-0.3473	-0.0990	-0.1033	1.7056	0.4189	1.1123	0.1110
8	CH=C	-0.4408	0.2765	-0.1353	-0.0375	-0.3035	-0.1969	-0.5763	0.0312	0.0683	0.2096	0.2482	-0.0689
9	C=C	-0.3975	-0.0285	-0.9184	-0.4819	-0.8157	-0.4467	-1.4670	-0.2164	-0.4700	-0.0352	-0.5471	-0.1193
10	CH_2=C=CH	****	****	****	****	****	****	****	****	****	****	****	****
11	CH_2=C=C	****	****	****	****	****	****	****	****	****	****	****	****
12	C=C=C	****	****	****	****	****	****	****	****	****	****	****	****
13	CH≡C	-1.4524	-0.7521	-0.6233	-0.7808	-0.6030	-0.7958	-0.0924	0.1487	-0.5263	-0.4364	-0.6330	-0.5975
14	C≡C	****	****	****	****	****	****	****	****	****	****	****	****
15	aCH	0.1490	-0.0163	0.1566	0.0107	0.1167	0.0106	0.0640	-0.0121	0.1571	0.0233	0.1393	0.0255
16	aC fused with aromatic ring	-0.3459	0.0171	-0.4712	0.0209	-0.3477	-0.0973	-0.4930	-0.2515	-0.1678	0.2473	-0.2880	0.1559
17	aC fused with non-aromatic ring	-0.2140	-0.0898	-0.2943	-0.1356	-0.2763	-0.1125	-0.5172	-0.3238	-0.0410	0.2553	-0.1375	0.1006
18	aC except as above	-0.8266	-0.0240	-0.8915	-0.3353	-0.9367	-0.3258	-1.0364	-0.9082	-0.4495	0.2073	-0.4613	-0.0179
19	aN in aromatic ring	0.1644	-0.3598	-0.0029	-0.3664	0.2228	-0.1959	0.3106	0.1521	0.2173	-0.4254	0.2630	-0.3613
20	aC-CH_3	-0.2600	-0.1517	0.1183	0.0806	-0.1506	0.0707	-0.2347	-0.1264	-0.0250	0.2765	-0.1506	0.1845
21	aC-CH_2	-0.2461	-0.0119	-0.5533	-0.0096	-0.3687	-0.0487	-0.4396	-0.0241	-0.1477	0.6430	-0.1678	0.3112
22	aC-CH	-0.0544	-0.2853	-0.5470	-0.5044	-0.6417	-0.7429	-0.0676	-0.3662	-0.0490	-0.3446	0.0021	-0.4457
23	aC-C	-0.6955	-0.4066	-1.1136	-0.7185	-1.2866	-1.3071	-1.6503	-0.9373	-0.8133	-0.5220	-0.8327	-0.5737
24	aC-CH=CH_2	-0.0279	-0.0406	0.4409	0.8714	1.6304	1.6651	0.1127	0.8952	0.8303	1.6498	1.1118	1.6201
25	aC-CH=CH	0.3962	-0.6772	-0.0339	-1.2093	-0.0013	-0.9732	-0.5925	-2.3205	0.8344	-0.5988	0.4860	-0.6970
26	aC-C=CH_2	****	****	****	****	****	****	****	****	****	****	****	****
27	aC-C≡CH	****	****	****	****	****	****	****	****	****	****	****	****
28	aC-C≡C	****	****	****	****	****	****	****	****	****	****	****	****
29	OH	0.1973	-0.1329	0.1224	-0.0335	0.3576	0.1939	0.5909	0.4026	0.0373	-0.3806	0.1133	-0.2356
30	aC-OH	0.3579	-0.2124	0.2272	-0.3507	0.4162	-0.1482	0.6553	0.4879	0.6292	0.3186	0.6065	0.1571
31	COOH	0.4136	-0.4133	0.1754	-0.7440	0.4426	0.0566	0.8145	1.3425	0.1276	-0.4999	0.1350	-0.6773
32	aC-COOH	-0.4401	0.0701	0.3832	-0.2474	0.4901	0.2991	1.1441	1.6362	0.6152	0.0521	0.3617	-0.0445
33	CH_3CO	-0.5659	0.2394	-1.0870	0.4385	-0.7815	0.8784	-0.8496	0.7182	-0.5576	0.9002	-0.7829	0.7620
34	CH_2CO	-1.2892	****	-1.4790	****	-0.6787	****	0.1067	****	-2.2429	****	-1.9003	****
35	CHCO	****	-0.4946	****	-0.6214	****	-0.2234	****	-0.0715	****	-0.2631	****	-0.5728
36	CCO	****	****	****	****	****	****	****	****	****	****	****	****
37	aC-CO	-0.6823	-0.2301	-1.7511	-0.8199	-0.8861	-0.8717	-0.2081	-0.7125	-1.3882	0.0514	-1.3193	-0.5369
38	CHO	0.9768	-1.2611	0.6049	-0.8794	0.7182	-0.4964	0.7156	-0.0583	0.5516	-0.8968	0.5158	-0.9705
39	aC-CHO	1.4448	-0.1404	1.3160	-0.1508	1.3270	-0.1747	1.2618	-0.2802	1.0550	0.4847	0.9880	0.0184

40	CH_3COO	0.5459	0.9003	0.4241	0.6052	0.7444	1.4409	0.3646	1.5768	0.2724	0.9531	0.2505	0.8098
41	CH_2COO	0.6057	1.0128	0.0881	0.7938	0.5648	1.5247	0.9133	2.4827	-0.3351	1.4263	0.1243	1.3393
42	CHCOO	****	-0.0308	****	-0.4444	****	1.0059	****	0.4328	****	1.3017	****	0.2341
43	CCOO	-0.8771	****	-1.1997	****	-1.4991	****	-1.2924	****	-1.7785	****	-1.8290	****
44	HCOO	****	****	****	****	****	****	****	****	****	****	****	****
45	aC-COO	-0.3185	0.2823	-0.4614	-0.0342	0.0421	0.3270	-0.0685	0.7594	0.2976	0.5649	0.0071	0.2475
46	aC-OOCH	****	****	****	****	****	****	****	****	****	****	****	****
47	aC-OOC	0.6800	0.8397	1.1171	0.2293	4.3179	0.6441	7.4590	0.3575	3.2164	1.1469	1.2900	0.6910
48	COO except as above	-0.1330	0.0421	-0.5584	-0.1768	-0.0621	0.3225	0.5514	0.7978	-0.2484	0.6865	-0.1042	0.3132
49	CH_3O	-0.0737	0.3926	0.0890	0.3819	0.3565	0.2597	-0.0584	0.3986	-0.0537	-0.1449	0.1894	0.0078
50	CH_2O	0.5146	0.3708	0.5731	0.1905	0.2700	0.2248	0.1110	0.2236	0.2602	-0.2207	0.1187	-0.1002
51	CH-O	****	0.2217	****	0.2546	****	-0.1904	****	0.8482	****	0.2456	****	-0.1258
52	C-O	0.7438	****	0.3865	****	0.9454	****	2.4004	****	0.9286	****	0.1394	****
53	aC-O	-0.1124	-0.2269	-0.1606	-0.3367	-0.1500	-0.4353	0.1490	-0.1858	-0.0825	-0.1939	-0.1134	-0.1872
54	CH_2NH_2	2.6417	****	2.2328	****	2.3637	****	2.6710	****	1.9022	****	2.1752	****
55	$CHNH_2$	-0.4403	****	-0.4944	****	-0.6307	****	0.2293	****	-0.9306	****	-0.6118	****
56	CNH_2	****	****	****	****	****	****	****	****	****	****	****	****
57	CH_3NH	0.5989	1.1599	1.2116	0.8850	0.1976	1.7821	0.3662	1.8835	0.2348	2.1442	0.2376	1.6882
58	CH_2NH	-0.2935	2.6870	0.2179	2.9085	-0.4139	1.6683	-0.3472	0.4786	-0.3451	2.2910	-0.2829	2.6186
59	CHNH	****	-1.0359	****	-0.8527	****	-1.5413	****	-0.8592	****	-1.9285	****	-1.9808
60	CH_3N	-0.9628	0.5031	-0.7580	-0.0732	-0.8942	0.0312	-0.7265	0.5274	-0.7908	0.1335	-0.7504	0.2359
61	CH_2N	-0.0865	-2.9995	-0.1413	-2.5338	-0.3720	-3.8875	-0.1841	-3.2261	-0.2994	-3.0576	-0.1682	-2.5588
62	aC-NH_2	0.3571	-0.2153	0.2748	-0.3095	0.1398	-0.2644	0.4196	-0.0304	0.2464	-0.1121	0.2767	-0.1759
63	aC-NH	-0.0048	0.2582	-0.2628	-0.1291	-0.2814	-0.2055	-0.1881	-0.0031	-0.2280	0.1322	-0.2492	0.0722
64	aC-N	-0.4604	1.0680	-0.4588	0.7264	-0.9315	0.3995	-0.9075	0.1869	-0.8443	0.5837	-0.9327	0.2939
65	NH_2 except as above	-0.1142	0.0985	-0.4335	0.0042	-0.2549	0.2402	0.0728	1.3282	-0.3473	0.4343	-0.3593	0.2852
66	CH=N	0.0044	-0.4784	0.3889	-0.7022	0.3618	-0.9722	0.3871	-0.5138	0.7647	-0.5477	0.8108	-0.7022
67	C=N	0.3823	0.1238	-0.0164	0.0187	0.2669	-0.6474	0.1658	-0.0183	1.0080	0.5949	0.9758	0.3875
68	CH_2CN	****	****	****	****	****	****	****	****	****	****	****	****
69	CHCN	****	****	****	****	****	****	****	****	****	****	****	****
70	CCN	-1.8216	****	-1.9527	****	-1.7907	****	-1.5534	****	-2.2276	****	-2.1245	****
71	aC-CN	1.2567	-0.1886	0.9475	-0.2222	1.2346	-0.0152	1.5580	0.2173	1.3322	-0.1325	1.3219	0.0023
72	CN except as above	0.0264	-0.1436	-0.7595	-0.1718	-0.1286	0.3670	0.0422	0.0000	-0.1534	0.0266	-0.2672	-0.0738
73	CH_2NCO	****	****	****	****	****	****	****	****	****	****	****	****
74	CHNCO	****	****	****	****	****	****	****	****	****	****	****	****
75	CNCO	****	****	****	****	****	****	****	****	****	****	****	****
76	aC-NCO	1.1089	-1.9959	0.5349	-2.1334	0.4265	-2.0027	-0.2182	-2.4080	0.9661	-1.5508	0.8283	-1.7168
77	CH2NO_2	****	****	****	****	****	****	****	****	****	****	****	****
78	$CHNO_2$	-0.9658	-0.9958	-1.7870	-1.4643	-1.0063	-0.6800	-2.0912	-0.9445	-1.5018	-0.5350	-1.1962	-0.6291
79	CNO_2	0.7609	****	1.1691	****	0.4194	****	-1.2488	****	0.7154	****	0.7829	****
80	aC-NO_2	0.0987	-0.7413	-0.5705	-0.9110	-0.1609	-0.6097	-0.1511	-0.4803	-0.0347	-0.3346	0.0709	-0.4266
81	NO_2 except as above	-0.6643	0.2770	-1.0546	0.2802	-0.6147	0.5173	-0.5777	0.5233	-0.7462	1.3837	-0.7208	1.1761
82	ONO	0.6316	****	-0.0875	****	0.2952	****	-0.7739	****	-0.2319	****	0.2707	****
83	ONO_2	0.4802	****	0.0600	****	0.2000	****	0.3815	****	0.0576	****	0.1529	****
84	$HCON(CH_2)_2$	****	****	****	****	****	****	****	****	****	****	****	****
85	$HCONHCH_2$	****	****	****	****	****	****	****	****	****	****	****	****

86	$CONH_2$	0.0134	-0.4139	0.1691	-0.0586	0.8177	-0.1241	1.4121	0.8899	0.7314	-0.5006	0.8945	0.0155
87	$CONHCH_3$	-0.0313	-0.8362	-0.7594	-1.0911	-0.4951	-0.8312	-0.1024	0.2246	-0.8420	-0.8833	-0.5860	-1.0145
88	$CONHCH_2$	1.9518	1.1442	1.4832	1.4330	1.2426	1.1152	1.5296	1.6594	2.2234	2.1503	2.1266	1.9013
89	$CON(CH_3)_2$	-0.6537	-0.0950	-1.0803	-0.3867	-1.1343	-0.1322	-0.1937	0.6511	-1.3163	-0.2674	-1.2917	-0.3364
90	$CONCH_3CH_2$	****	****	****	****	****	****	****	****	****	****	****	****
91	$CON(CH_2)_2$	-0.8273	-0.3936	-0.9693	-0.7407	-1.3969	-0.8424	-1.1112	-0.1034	-1.1740	-0.5455	-0.9545	-0.8783
92	CONHCO	****	****	****	****	****	****	****	****	****	****	****	****
93	CONCO	****	-0.9068	****	-1.4827	****	-0.8258	****	0.3186	****	-1.0391	****	-0.8294
94	aC-$CONH_2$	0.8871	****	-0.0865	****	0.8239	****	0.5860	****	0.2734	****	0.4742	****
95	aC-NH(CO)H	****	****	****	****	****	****	****	****	****	****	****	****
96	aC-N(CO)H	****	****	****	****	****	****	****	****	****	****	****	****
97	aC-CONH	-0.7583	1.2697	-1.4627	0.6663	-0.6019	1.0400	-0.2033	2.0214	-1.3456	1.0294	-0.8983	0.6978
98	aC-NHCO	-0.0004	-0.0637	-0.2983	-0.5320	-0.1121	0.0350	0.3804	0.1428	-0.4114	0.2488	-0.2086	0.1182
99	aC-(N)CO	-1.0967	0.3045	-1.4759	-0.3026	-1.3678	-0.4157	-0.5218	0.6672	-0.6634	-0.2291	-0.4365	-0.2363
100	NHCONH	1.4303	****	0.9832	****	1.1946	****	1.9429	****	0.4362	****	0.5252	****
101	NH_2CONH	0.3821	****	-0.3543	****	0.0467	****	0.7847	****	-0.0753	****	0.0593	****
102	NH_2CON	-0.4982	****	-1.1906	****	-0.7382	****	-0.0828	****	-0.7912	****	-0.7436	****
103	NHCON	0.3024	0.0100	-0.5509	-0.7257	0.2361	0.2184	0.6711	1.3846	0.3913	-0.6951	0.2303	-0.4350
104	NCON	****	****	****	****	****	****	****	****	****	****	****	****
105	aC-$NHCONH_2$	****	****	****	****	****	****	****	****	****	****	****	****
106	aC-NHCONH	1.4478	1.0207	1.2121	0.5931	0.9761	0.4899	1.8983	0.5083	1.4380	-0.0717	1.3208	0.0327
107	NHCO except as above	0.1199	-0.0146	0.0454	0.0032	0.2878	0.2213	0.8196	0.3913	0.2920	0.0295	0.3624	0.0287
108	CH_2Cl	-0.3830	-0.7636	-0.3819	-0.6860	-0.3472	-0.3273	-0.4563	-0.8939	-0.3392	-0.6522	-0.4056	-0.5439
109	CHCl	0.6637	0.1485	-0.2285	-0.2019	0.5142	-0.6331	-0.9817	-0.9698	0.1789	-0.2755	0.2965	0.1149
110	CCl	****	****	****	****	****	****	****	****	****	****	****	****
111	$CHCl_2$	0.4531	-0.7610	-0.1357	-1.1547	0.3226	-0.5988	-0.6482	-1.1267	0.0549	-0.6492	0.0719	-0.6546
112	CCl_2	****	0.0162	****	-0.1154	****	-0.4194	****	-1.0026	****	0.1964	****	0.0306
113	CCl_3	-0.0116	-1.2569	-0.6886	-1.3607	-0.3901	-1.2703	-1.5385	-1.5906	-0.6039	-0.9174	-0.5410	-0.9872
114	CH_2F	1.8052	1.9109	1.2625	1.2313	1.6943	2.6100	1.0269	1.5339	1.2637	2.2373	1.3172	1.8204
115	CHF	****	****	****	****	****	****	****	****	****	****	****	****
116	CF	****	****	****	****	****	****	****	****	****	****	****	****
117	CHF_2	****	-0.0027	****	0.1968	****	0.2337	****	-0.6164	****	0.8043	****	0.3066
118	CF_2	****	****	****	****	****	****	****	****	****	****	****	****
119	CF_3	0.7993	-0.4068	0.0981	-0.4311	0.5850	-0.6634	-0.8071	-1.1163	0.1960	-0.8828	0.3345	-0.4122
120	CCl_2F	****	-1.4244	****	-2.2499	****	-1.2401	****	-2.6623	****	-1.6824	****	-1.6484
121	HCClF	****	1.3816	****	0.4334	****	1.5902	****	0.3557	****	1.1147	****	1.1246
122	$CClF_2$	****	0.9918	****	1.0841	****	1.2244	****	1.0152	****	1.0564	****	1.0765
123	aC-Cl	-0.1496	-0.2926	-0.2917	-0.4226	-0.3034	-0.3287	-0.7378	-0.7018	-0.3028	-0.2349	-0.2799	-0.2913
124	aC-F	-0.5147	-0.6887	-1.2858	-0.5469	-0.8683	-0.4136	-1.0689	-0.4146	-0.8543	-0.2960	-1.0622	-0.3001
125	aC-I	****	****	****	****	****	****	****	****	****	****	****	****
126	aC-Br	0.1178	-0.2363	-0.0015	-0.2642	0.1150	-0.1578	-0.0861	-0.2954	0.1426	-0.1230	0.0672	-0.1986
127	-I except as above	****	****	****	****	****	****	****	****	****	****	****	****
128	-Br except as above	-0.0869	-0.1776	-0.0596	-0.2436	-0.2059	0.2137	-0.3641	-0.1295	-0.3791	-0.1064	-0.2836	-0.0557
129	-F except as above	0.3896	-0.1807	0.3099	0.0680	0.4146	0.1057	0.3737	0.0762	0.2175	-0.0558	0.2341	-0.0119
130	-Cl except as above	0.0452	-0.2286	-0.0167	-0.2860	-0.0168	-0.0381	-0.1609	-0.3444	-0.0951	-0.2142	-0.0896	-0.2182
131	CHNOH	****	****	****	****	****	****	****	****	****	****	****	****

132	CNOH	0.0478	****	-0.7777	****	-0.3279	****	0.1303	****	-0.1868	****	0.0328	****
133	aC-CHNOH	****	****	****	****	****	****	****	****	****	****	****	****
134	OCH_2CH_2OH	2.1707	-0.1215	1.7390	-0.2071	1.8167	-0.0192	2.7585	1.5986	1.3109	-0.1664	1.3937	-0.2047
135	$OCHCH_2OH$	****	****	****	****	****	****	****	****	****	****	****	****
136	OCH_2CHOH	****	1.0611	****	1.0028	****	0.9576	****	2.7896	****	1.1904	****	1.0355
137	-O-OH	****	****	****	****	****	****	****	****	****	****	****	****
138	CH_2SH	-0.7931	****	-1.8639	****	-1.3399	****	-0.5572	****	-0.2854	****	-0.2384	****
139	CHSH	****	****	****	****	****	****	****	****	****	****	****	****
140	CSH	****	****	****	****	****	****	****	****	****	****	****	****
141	aC-SH	-1.6674	****	-3.2165	****	-2.0555	****	-0.7549	****	-2.3760	****	-2.0847	****
142	-SH except as above	****	****	****	****	****	****	****	****	****	****	****	****
143	CH_3S	1.7560	-1.1500	-0.1111	-0.4163	0.0458	-0.3930	0.4226	-1.2235	0.1624	-1.4553	-0.0226	-1.2049
144	CH_2S	-0.3952	-0.2160	-0.9617	-0.2021	-1.1017	-0.3669	-1.3556	-0.6114	-0.9489	-0.2602	-0.9053	-0.4426
145	CHS	****	-0.1118	****	-0.3187	****	-1.2181	****	-1.6430	****	-1.7585	****	-1.3933
146	CS	1.6458	-0.0620	1.4751	0.5075	1.2477	-1.0085	0.4709	-0.6864	1.0890	-0.8233	1.1165	-0.8747
147	aC-S-	-1.1541	-0.2388	-1.6190	-0.6546	-1.3791	-0.6194	-1.5459	-0.9009	-1.1808	-0.2158	-1.2017	-0.4052
148	SO	1.4991	1.2728	0.5711	1.4900	0.2190	1.3598	0.4583	0.6590	0.2160	1.7927	0.4250	1.1434
149	SO2	****	-0.7598	****	-1.1765	****	-0.3042	****	-0.4479	****	-0.0085	****	-0.0109
150	SO_3 (sulfite)	0.6697	0.6069	-0.1826	0.1647	-0.5130	0.3638	-0.8241	-0.7865	-0.2758	0.2651	0.0493	0.1334
151	SO_3 (Sulfonate)	-0.3168	****	-1.3506	****	-0.5966	****	0.4606	****	-0.8646	****	-0.6894	****
152	SO_4 (Sulfate)	****	****	****	****	****	****	****	****	****	****	****	****
153	aC-SO	****	-1.5062	****	-1.6110	****	-1.3267	****	-0.0555	****	-1.4480	****	-1.2697
154	aC-SO_2	0.2941	-0.6688	-0.4563	-1.0517	0.0824	-0.7098	0.5214	-0.4493	-0.2997	0.0352	-0.1496	-0.2397
155	PH (phosphine)	****	****	****	****	****	****	****	****	****	****	****	****
156	P (Phospine)	****	-1.2820	****	-1.5965	****	-2.3001	****	-1.6462	****	-0.6856	****	-0.0428
157	PO_3 (Phospite)	****	****	****	****	****	****	****	****	****	****	****	****
158	PHO_3 (Phosponate)	0.7416	****	-0.0104	****	0.3076	****	1.1965	****	0.0906	****	0.2658	****
159	PO_3 (Phosponate)	1.8953	-0.4482	1.3651	-1.2315	1.2858	-0.4895	2.0993	1.0570	1.4417	-0.6521	1.5191	-0.1747
160	PHO_4 (Phospate)	****	****	****	****	****	****	****	****	****	****	****	****
161	PO_4 (Phospate)	1.0798	-2.2565	-0.2379	-2.1906	0.7802	-2.1982	1.4306	-0.5733	0.9302	-2.1455	0.6254	-1.8969
162	aC-PO_4	****	****	****	****	****	****	****	****	****	****	****	****
163	aC-P	****	****	****	****	****	****	****	****	****	****	****	****
164	CO_3 (Carbonate)	****	****	****	****	****	****	****	****	****	****	****	****
165	C_2H_3O	0.4798	-1.4409	-0.0067	-1.6703	0.4151	-1.0459	-0.2880	-1.2669	-0.3115	-1.0362	0.0989	-1.1662
166	C_2H_2O	****	-0.2548	****	-0.2048	****	-0.4602	****	-0.3366	****	-0.9231	****	-0.9071
167	C_2HO	****	****	****	****	****	****	****	****	****	****	****	****
168	CH_2 (cyclic)	0.0517	0.0477	0.0258	0.0949	-0.0231	0.0852	-0.0035	0.0641	0.1168	0.1538	0.0651	0.1094
169	CH (cyclic)	-0.2099	-0.1818	-0.3139	-0.2209	-0.2499	-0.3282	-0.4216	-0.3743	-0.1332	-0.0170	-0.1400	-0.0728
170	C (cyclic)	-0.2783	-0.1269	-0.3415	-0.1110	-0.3659	-0.3299	-0.4277	-0.1314	-0.1758	-0.0743	-0.1693	-0.0609
171	CH=CH (cyclic)	0.2143	-0.3422	0.1756	-0.2311	-0.2314	-0.2424	-0.1734	-0.5072	0.2430	0.2965	0.1448	0.1844
172	CH=C (cyclic)	0.3631	0.1375	0.2563	0.1158	-0.0462	0.0647	-0.0189	-0.0488	0.4705	0.3370	0.2823	0.1311
173	C=C (cyclic)	-0.1056	-0.7511	-0.2382	-0.9559	-0.4197	-1.3426	-0.3613	-1.0790	0.1772	-0.9770	0.1060	-0.8705
174	CH_2=C (cyclic)	-0.1864	****	-0.5550	****	-1.2561	****	-1.4313	****	0.0896	****	0.3598	****
175	NH (cyclic)	0.0173	-0.2067	-0.0522	-0.3755	0.0650	-0.2712	0.2082	0.2561	0.0309	-0.5038	0.0956	-0.3555
176	N (cyclic)	0.0067	-0.3155	-0.1687	-0.4413	-0.0415	-0.5931	0.2860	-0.1578	-0.3214	-1.3094	-0.1979	-1.0750
177	CH=N (cyclic)	0.8193	0.1688	0.6106	0.1964	0.6874	0.3340	0.9412	0.2484	0.3549	0.6275	0.3681	0.7590

178	C=N (cyclic)	-0.2137	-0.3953	-0.3968	-0.4652	-0.1178	-0.3734	0.1774	-0.1057	-0.1374	0.0737	0.0007	0.0397
179	O (cyclic)	-0.5237	-0.1128	-0.3063	-0.1283	-0.2134	-0.1151	-0.1357	0.0910	-0.7957	-0.5423	-0.6411	-0.4295
180	CO (cyclic)	0.0827	0.2883	-0.0578	0.1261	0.1300	0.4942	0.4265	0.6506	0.0014	0.6848	0.0099	0.5919
181	S (cyclic)	-0.2378	0.0549	-0.2629	0.3156	-0.1562	0.0932	-0.2200	0.3792	0.0032	0.0763	0.0118	-0.0247
182	SO_2 (cyclic)	0.1035	-0.1536	-0.7377	-0.4337	0.9810	-0.2873	1.0942	0.1567	3.4223	-0.5213	3.3165	-0.5782
183	>NH	-0.4571	-0.0645	-0.3987	-0.3324	-0.2798	-0.3977	-0.5169	-0.7340	-0.2534	-0.7876	-0.2042	-0.6020
184	-O-	1.7348	0.1749	0.3176	-0.5076	0.1044	0.5573	-0.6604	0.6133	0.2070	0.5392	-0.6565	0.6999
185	-S-	1.5902	0.3574	1.3400	0.3992	0.9095	0.2659	0.5240	0.1034	1.7518	-0.2023	1.6974	-0.0040
186	>CO	0.0857	-0.5621	****	-0.8385	-0.6724	-0.4349	-0.5629	-0.2143	-0.2026	-0.3192	****	-0.4260
187	PO_2	****	****	****	****	****	****	****	****	****	****	****	****
188	CH-N	****	0.2157	-0.7919	0.2827	****	-0.7992	****	0.1344	****	0.2756	-0.6340	0.0122
189	SiHO	****	****	****	****	****	****	****	****	****	****	****	****
190	SiO	****	****	****	****	****	****	****	****	****	****	****	****
191	SiH_2	****	****	****	****	****	****	****	****	****	****	****	****
192	SiH	****	****	****	****	****	****	****	****	****	****	****	****
193	Si	****	****	****	****	****	****	****	****	****	****	****	****
194	$(CH_3)_3N$	****	****	****	****	****	****	****	****	****	****	****	****
195	N=N	****	0.0102	****	-0.1583	****	0.0432	****	0.1089	****	-0.4251	****	-0.1029
196	C_{cyclic}=N-	-0.5479	-0.8507	-0.7442	-0.7851	-0.1413	-0.3556	-0.6733	-0.2041	0.3919	0.9717	0.1476	0.5806
197	C_{cyclic}=CH-	-0.3670	****	-0.6552	****	-0.5085	****	-0.6474	****	-0.9219	****	-0.7721	****
198	C_{cyclic}=NH	****	****	****	****	****	****	****	****	****	****	****	****
199	N=O	-0.9060	****	-0.8677	****	-0.9591	****	-0.6917	****	-1.1569	****	-1.0287	****
200	C_{cyclic}=C	-0.7306	****	-1.0157	****	-1.0434	****	-1.0337	****	-1.0385	****	-0.9418	****
201	P=O	-0.6489	-1.7203	-0.0055	-1.9083	-0.0211	-1.6554	0.6175	-0.5475	-0.1279	-0.8717	-0.0973	-1.3289
202	N=N	0.3441	****	-0.6377	****	-0.1406	****	-0.6545	****	-0.6320	****	-0.6844	****
203	C=NH	****	****	****	****	****	****	****	****	****	****	****	****
204	>C=S	0.4296	-1.0799	0.1655	-0.5831	0.0996	-0.8105	0.5208	-0.5155	-0.0525	0.0162	-0.0481	-0.3127
205	aC-CON	-0.2883	-1.6096	-0.9420	-2.0087	-0.7021	-2.6680	0.4358	-3.0736	-0.4751	-2.7795	-0.4727	-1.9244
206	aC=O	2.0541	****	2.0441	****	2.1876	****	2.6538	****	2.4073	****	2.2166	****
207	aN-	-0.7235	-0.8058	-0.5085	-1.0327	-0.5252	-0.7093	-0.0804	0.2873	-0.0185	-0.0801	0.3597	-0.0003
208	-Na	****	****	****	****	****	****	****	****	****	****	****	****
209	-K	****	****	****	****	****	****	****	****	****	****	****	****
210	HCONH	0.6315	****	0.1483	****	0.5060	****	1.0373	****	0.1419	****	0.1964	****
211	CHOCH	****	****	****	****	****	****	****	****	****	****	****	****
212	C_2O	****	****	****	****	****	****	****	****	****	****	****	****
213	SiH_3	****	****	****	****	****	****	****	****	****	****	****	****
214	SiH_2O	****	****	****	****	****	****	****	****	****	****	****	****
215	CH=C=CH	****	****	****	****	****	****	****	****	****	****	****	****
216	CH=C=C	****	****	****	****	****	****	****	****	****	****	****	****
217	OP(=S)O	0.9193	-1.0765	0.8755	-1.0305	-0.0055	-1.6725	1.4290	-1.1229	0.1428	-0.8306	-0.2776	-1.1215
218	R	****	****	****	****	****	****	****	****	****	****	****	****
219	$CF_{2cyclic}$	****	****	****	****	****	****	****	****	****	****	****	****
220	CF_{cyclic}	-1.1221	****	-1.5038	****	-1.3502	****	-2.3291	****	-0.9893	****	-1.0702	****

[a] The symbols $EUA_{C\,1i}$, $EUA_{NC\,1i}$, $ERA_{C\,1i}$, $ERA_{NC\,1i}$, $EFW_{C\,1i}$, $EFW_{NC\,1i}$, $ESW_{C\,1i}$, $ESW_{NC\,1i}$, $ENS_{C\,1i}$, $ENS_{NC\,1i}$, $EAS_{C\,1i}$, and $EAS_{NC\,1i}$ represent the contributions (C_i) of the first-order groups for the corresponding properties.

Table B3. MG Method Based Property Models Analysed Using Step-Wise Regression Method: Second-Order Groups and their Contributions [a] for the Properties÷ LC_{50}(FM), LC_{50}(DM), LD_{50}, $LogW_s$, BCF, PEL(OSHA-TWA), and PCO

	Group	$LC50(FM)_{2j}$	$LC50(DM)_{2j}$	$LD50_{2j}$	$LogWs_{2j}$	BCF_{2j}	PEL_{2j}	PCO_{2j}
1	$(CH_3)_2CH$	-0.0764	-0.1057	-0.0308	-0.0961	0.1168	-0.0213	0.0154
2	$(CH_3)_3C$	0.0412	0.0599	0.0840	-0.3055	-0.0811	0.0227	0.0272
3	$CH(CH_3)CH(CH_3)$	-0.1113	****	0.0468	-0.1821	1.0401	0.0043	0.0580
4	$CH(CH_3)C(CH_3)_2$	****	****	-0.1376	0.4430	0.8297	0.0346	-0.0299
5	$C(CH_3)_2C(CH_3)_2$	****	****	0.0640	-0.5905	****	-0.0239	0.0148
6	CH_n=CH_m-CH_p=CH_k (*k,m,n,p* in 0..2)	-0.3839	0.4580	0.0901	-0.7096	-0.5140	0.2521	0.3386
7	CH_3-CH_m=CH_n (*m,n* in 0..2)	0.2163	-1.5389	-0.0529	-0.0474	0.2127	0.3417	-0.0467
8	CH_2-CH_m=CH_n (*m,n* in 0..2)	-0.2321	0.1522	-0.0234	-0.0205	0.0272	0.2745	-0.0203
9	CH_p-CH_m=CH_n (*m,n* in 0..2; *p* in 0..1)	0.2029	-0.3690	0.1804	0.1311	-0.1108	****	-0.0588
10	CHCHO or CCHO	0.1432	****	0.0030	-0.3170	****	****	-0.1355
11	CH_3COCH_2	-0.0733	0.2120	-0.0806	0.1427	****	-0.0934	-0.0677
12	CH_3COCH or CH_3COC	0.4132	****	0.2043	0.5001	0.0430	-0.1821	0.0025
13	CHCOOH or CCOOH	-0.6512	****	0.1233	0.3928	-0.2597	0.5903	0.0367
14	CH_3COOCH or CH_3COOC	-0.0341	****	0.0951	0.1888	****	-0.1284	-0.0162
15	CO-O-CO	****	****	-0.0807	-0.6322	0.0000	1.0520	****
16	CHOH	0.0554	1.0759	0.0219	0.4393	-0.2086	0.0954	0.0132
17	COH	0.0625	****	0.0891	0.7215	-0.1750	-0.1548	0.3195
18	CH_3COCH_nOH (*n* in 0..2)	****	****	-0.0982	0.8644	****	****	-0.1783
19	NCCHOH or NCCOH	****	****	0.1495	0.6017	****	0.0000	****
20	OH-CH_n-COO (*n* in 0..2)	0.8775	****	-0.0409	0.2457	-0.4364	****	-0.2612
21	$CH_m(OH)CH_n(OH)$ (*m,n* in 0..2)	-1.2012	-1.1768	0.0066	-0.1728	-0.5661	0.2618	-0.0693
22	$CH_m(OH)CH_n(NH_p)$ (*m,n,p* in 0..2)	-0.1316	0.6018	-0.0205	0.2429	0.0000	0.3565	-0.0042
23	$CH_m(NH_2)CH_n(NH_2)$ (*m,n* in 0..2)	0.0265	0.0777	-0.4353	-0.7650	****	-0.4550	****
24	$CH_m(NH)CH_n(NH_2)$ (*m,n* in 1..2)	****	****	-0.1929	0.3688	0.0000	-0.3635	****
25	$H_2NCOCH_nCH_mCONH_2$ (*m,n* in 1..2)	****	****	****	-0.6827	****	****	****
26	$CH_m(NH_n)$-COOH (*m,n* in 0..2)	0.4575	0.6214	-0.1690	-0.1789	****	****	****
27	HOOC-CH_n-COOH (*n* in 1..2)	****	****	-0.1155	0.5780	****	****	****
28	HOOC-CH_n-CH_m-COOH (*n, m* in 1..2)	****	****	0.0445	0.2286	1.8872	****	-0.3516
29	HO-CH_n-COOH (*n* in 1..2)	****	****	-0.0777	-0.3265	0.2712	****	0.1355
30	NH_2-CH_n-CH_m-COOH (*n, m* in 1..2)	****	****	****	-0.6598	****	****	****
31	CH_3-O-CH_n-COOH (n in 1..2)	****	****	****	0.7930	****	****	****
32	HS-CH-COOH	0.0000	****	0.6687	0.2806	****	-1.2326	****
33	HS-CH_n-CH_m-COOH (*n, m* in 1..2)	****	****	0.0364	0.8204	****	****	****
34	NC-CH_n-CH_m-CN (*n, m* in 1..2)	****	****	0.5699	0.1470	****	-0.4976	****

35	OH-CH_n-CH_m-CN (*n, m* in 1..2)	****	****	0.2069	0.9963	****	****	****
36	HS-CH_n-CH_m-SH (*n, m* in 1..2)	****	****	-1.5941	****	****	****	****
37	COO-CH_n-CH_m-OOC (*n, m* in 1..2)	****	****	0.1041	-0.3202	****	****	-0.1422
38	OOC-CH_m-CH_m-COO (*n, m* in 1..2)	-0.1100	-0.1899	-0.0098	0.3351	0.0000	0.0000	0.3766
39	NC-CH_n-COO (*n* in 1..2)	****	****	-0.2553	0.0320	****	****	****
40	$COCH_n$COO (*n* in 1..2)	****	****	0.0455	0.2919	****	****	****
41	CH_m-O-CH_n=CH_p (*m,n,p* in 0..3)	****	****	-0.3996	-0.0117	1.1102	****	****
42	CH_m=CH_n-F (*m,n* in 0..2)	****	****	0.0100	-0.3430	****	-0.0351	****
43	CH_m=CH_n-Br (*m,n* in 0..2)	****	-0.0478	-0.3156	0.3983	****	0.8884	****
44	CH_m=CH_n-I (*m,n* in 0..2)	****	****	****	0.0121	****	****	****
45	CH_m=CH_n-Cl (*m,n* in 0..2)	0.0205	0.0144	0.1063	0.1214	0.0502	-0.1321	-0.0613
46	CH_m=CH_n-CN (*m,n* in 0..2)	-0.2490	0.4082	0.1708	1.0178	****	-0.1118	0.0000
47	CH_n=CH_m-COO-CH_p (*m,n,p* in 0..3)	0.0570	0.4720	-0.0032	0.2879	****	-0.2021	0.0143
48	CH_m=CH_n-CHO (*m,n* in 0..2)	1.3458	2.0090	0.2139	0.4115	****	0.1212	0.2420
49	CH_m=CH_n-COOH (*m,n* in 0..2)	1.0013	****	-0.0370	0.0592	****	-0.5149	-0.2860
50	aC-CH_n-X (*n* in 1..2) X: Halogen	-0.0077	****	-0.0723	-0.0254	-0.2510	0.2746	****
51	aC-CH_n-NH_m (*n* in 1..2; *m* in 0..2))	-0.1092	****	-0.1196	0.5674	0.0255	****	****
52	aC-CH_n-O- (*n* in 1..2)	0.3513	-0.1209	0.1138	0.5302	-0.0011	****	****
53	aC-CH_n-OH (*n* in 1..2)	-0.1727	****	-0.0304	0.1802	-0.0459	****	-0.0430
54	aC-CH_n-CN (*n* in 1..2)	1.4486	-0.1985	0.0300	0.5431	-0.7076	****	****
55	aC-CH_n-CHO (*n* in 1..2)	****	****	-0.2336	-0.2078	****	****	****
56	aC-CH_n-SH (*n* in 1..2)	****	****	0.2557	****	****	****	****
57	aC-CH_n-COOH (*n* in 1..2)	****	****	0.3508	0.1899	-0.5435	****	****
58	aC-CH_n-CO- (*n* in 1..2)	0.0000	****	0.0819	0.3811	-0.3243	****	****
59	aC-CH_n-S- (*n* in 1..2)	****	****	-0.4356	0.1455	-0.0417	****	****
60	aC-CH_n-OOC-H (*n* in 1..2)	****	****	0.1375	****	****	****	****
61	aC-CH_m-NO_2 (*n* in 1..2)	****	****	****	****	****	****	****
62	aC-CH_n-$CONH_2$ (*n* in 1..2)	****	****	0.2234	0.4008	****	****	****
63	aC-CH_n-OOC (*n* in 1..2)	-0.2284	-0.6714	-0.1266	-0.7736	-0.0125	0.5009	****
64	aC-CH_n-COO (*n* in 1..2)	-0.8038	0.5838	0.2030	-0.1084	0.4011	****	****
65	aC-SO_2-OH	-0.0528	****	-0.0221	0.6120	0.0638	****	****
66	aC-$CH(CH_3)_2$	-0.1479	-0.0398	-0.2498	-0.3480	0.0743	0.3027	0.0372
67	aC-$C(CH_3)_3$	0.4288	****	0.0038	-0.0489	-0.0393	0.4519	0.0160
68	aC-CF_3	-0.0138	-0.0943	-0.0622	0.0992	0.0751	****	0.0111
69	$(CH_n$=$C)_{cyc}$-CHO (*n* in 0..2)	****	****	-0.1006	-0.3013	****	-0.5824	****
70	$(CH_n$=$C)_{cyc}$-COO-CH_m (*n,m* in 0..3)	****	****	-0.0304	0.2300	-0.0373	****	****
71	$(CH_n$=$C)_{cyc}$-CO- (*n* in 0..2)	-0.5436	****	0.0327	-0.2010	****	****	****
72	$(CH_n$=$C)_{cyc}$-CH_3 (*n* in 0..2)	-0.1629	0.2697	-0.0287	0.1151	-0.1004	-0.2472	-0.1472
73	$(CH_n$=$C)_{cyc}$-CH_2 (*n* in 0..2)	0.0686	-2.1549	0.0605	0.4049	****	-0.6411	****
74	$(CH_n$=$C)_{cyc}$-CN (*n* in 0..2)	-0.6072	****	-0.6863	-0.2613	****	****	****

75	$(CH_n{=}C)_{cye}$-Cl (n in 0..2)	0.1507	0.0423	-0.0377	0.0344	0.0199	0.1901	****
76	CH_{cye}-CH_3	-0.0418	0.0679	0.0533	0.0404	-0.1411	-0.2157	-0.0279
77	CH_{cye}-CH_2	-0.5577	-1.2746	0.0237	0.1272	0.2565	-0.2110	-0.0028
78	CH_{cye}-CH	0.7411	-2.6422	-0.0129	0.0975	-0.2238	****	0.0378
79	CH_{cye}-C	-0.4529	****	-0.2425	-0.2294	0.4370	****	-0.1253
80	CH_{cye}-$CH{=}CH_n$ (n in 1..2)	0.3363	-2.3849	0.1087	-0.3632	-0.1528	1.2756	****
81	CH_{cye}-$C{=}CH_n$ (n in 1..2)	-0.7578	2.1395	-0.1690	0.0967	-1.1805	****	0.6709
82	CH_{cye}-Cl	-0.1520	-0.0695	0.1234	0.0072	-0.0429	-1.1409	****
83	CH_{cye}-F	****	****	****	-0.0294	****	****	****
84	CH_{cye}-OH	-0.1258	0.6398	0.0353	-0.1178	-0.3225	-0.3254	-0.0604
85	CH_{cye}-NH_2	****	****	-0.1432	0.1418	****	0.1139	****
86	CH_{cye}-NH-CH_n (n in 0..3)	0.2855	0.3464	-0.1079	0.4598	-0.0967	****	****
87	CH_{cye}-N-CH_n (n in 0..3)	0.7025	****	-0.1946	0.3802	0.4684	****	****
88	CH_{cye}-SH	****	****	1.4912	****	****	0.4203	****
89	CH_{cye}-CN	****	****	-0.3849	0.7628	****	****	****
90	CH_{cye}-COOH	-0.2236	****	0.0332	-0.1894	0.0867	****	****
91	CH_{cye}-CO	****	****	0.1850	0.1888	****	****	****
92	CH_{cye}-NO_2	****	****	0.5918	0.7474	****	****	****
93	CH_{cye}-S-	-1.1224	****	0.0123	0.6629	****	****	****
94	CH_{cye}-CHO	-0.0488	****	-0.1554	-0.1686	****	****	-0.0594
95	CH_{cye}-O-	****	****	-0.0867	0.0702	****	1.5492	****
96	CH_{cye}-OOCH	****	****	0.1650	****	****	****	****
97	CH_{cye}-COO	-0.4375	3.9708	-0.1197	-0.1930	0.5332	****	****
98	CH_{cye}-OOC	-0.9174	0.8222	-0.0947	0.2024	****	****	****
99	C_{cye}-CH_3	0.0618	-0.2693	-0.0245	-0.1054	0.0856	0.4295	0.0161
100	C_{cye}-CH_2	0.5211	0.8205	0.0520	0.2880	-0.1757	0.5001	-0.0554
101	C_{cye}-OH	-0.4611	0.9501	0.2298	-0.1743	-0.9768	-0.8443	****
102	$>N_{cye}$-CH_3	0.4824	0.2009	0.0145	0.1060	-0.3611	****	-0.1030
103	$>N_{cye}$-CH_2	-0.1675	0.0000	0.0030	0.2220	0.1704	****	****
104	AROMRINGs^{1}s^2	0.1461	0.1005	0.0332	-0.0987	0.0801	-0.0995	-0.0705
105	AROMRINGs^{1}s^3	-0.0819	0.0190	0.0470	-0.2251	0.0114	-0.2389	-0.1364
106	AROMRINGs^{1}s^4	0.0545	0.1187	-0.0160	-0.1804	-0.0626	0.1977	-0.0255
107	AROMRINGs^{1}s^{2}s^3	-0.1163	0.1842	0.1096	-0.0196	0.1842	0.2225	-0.1012
108	AROMRINGs^{1}s^{2}s^4	0.0807	-0.0351	-0.0055	-0.0269	-0.0430	-0.1142	0.0171
109	AROMRINGs^{1}s^{3}s^5	-0.1749	-0.1048	0.0383	-0.3208	0.4586	0.2071	-0.1189
110	AROMRINGs^{1}s^{2}s^{3}s^4	0.2641	-0.2547	-0.1754	0.0873	-0.0073	****	0.0249
111	AROMRINGs^{1}s^{2}s^{3}s^5	-0.1944	0.0897	-0.0836	0.0750	0.0998	-1.1173	0.1087
112	AROMRINGs^{1}s^{2}s^{4}s^5	0.1374	-0.1737	-0.0633	-0.1630	0.1481	****	0.0211
113	PYRIDINEs^2	-0.5596	****	0.3831	0.7714	-0.0366	****	****
114	PYRIDINEs^3	-0.1207	****	-0.0853	0.7486	-0.8944	****	****

115	PYRIDINEs^{4}	-0.3869	****	0.2497	0.3145	-0.2412	****	****
116	PYRIDINEs^{2}s^{3}	-0.8534	****	0.4131	0.3683	****	****	****
117	PYRIDINEs^{2}s^{4}	****	****	0.1771	0.5459	****	****	****
118	PYRIDINEs^{2}s^{5}	-0.2390	****	0.0415	0.1791	1.4201	****	****
119	PYRIDINEs^{2}s^{6}	0.0206	****	0.3342	-0.1606	****	****	****
120	PYRIDINEs^{3}s^{4}	****	****	0.2221	0.1017	****	****	****
121	PYRIDINEs^{3}s^{5}	****	****	-0.1887	-0.1816	****	****	****
122	PYRIDINEs^{2}s^{3}s^{6}	****	****	0.1607	-0.2204	****	****	****
123	(CH^{n}=CH^{m})^{cyc}-COOH	****	****	-0.8910	-0.4421	****	****	****
124	AROMRINGs^{1}s^{2}s^{3}s^{4}s^{5}	-0.1316	-0.4815	0.0886	0.3891	-0.0740	****	0.2418
125	aC-NHCOCH$_2$N	****	****	0.6234	1.2707	****	****	****
126	(N=C)$_{cyc}$-CH$_3$	-0.6390	****	0.0287	0.4988	-0.4374	****	****
127	aC-CONH(CH$_2$)$_2$N	****	****	0.3162	0.6654	****	****	****
128	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	-0.0116	-0.1908	-0.3573	****	****
129	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	0.3337	0.3478	****	****	****
130	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	-0.1481	-0.1134	-0.1719	****	****

[a] The symbols LC50(FM)$_{2j}$, LC50(DM)$_{2j}$LD50$_{2j}$, LogWs$_{2j}$, BCF$_{2j,}$ PEL$_{2j,}$ PCO$_{2j}$ represent the contributions (D_j) of the second-order groups for the corresponding properties. Note that there are no second-order contributions for GWP, ODP and AP since these properties belong to chemicals whose structure is defined by first-order groups alone.

Table B3 (continued). MG Method Based Property Models Analysed Using Step-Wise Regression Method: Second-Order Groups and their Contributions [a] for the Properties÷ EUA_C, EUA_{NC}, ERA_C, ERA_{NC}, EFW_C, EFW_{NC}, ESW_C, ESW_{NC}, ENS_C, ENS_{NC}, EAS_C, and EAS_{NC}.

	Group	$EUA_{C\,2j}$	$EUA_{NC\,2j}$	$ERA_{C\,2j}$	$ERA_{NC\,2j}$	$EFW_{C\,2j}$	$EFW_{NC\,2j}$	$ESW_{C\,2j}$	$ESW_{NC\,2j}$	$ENS_{C\,2j}$	$ENS_{NC\,2j}$	$EAS_{C\,2j}$	$EAS_{NC\,2j}$	$EUA_{C\,2j}$
1	$(CH_3)_2CH$	0.0493	0.1243	-0.0350	0.0690	-0.0267	0.1142	0.1385	0.1442	-0.0370	0.1454	-0.0149	0.1514	0.0493
2	$(CH_3)_3C$	-0.2156	-0.0108	0.0723	0.0970	-0.1280	-0.0997	-0.6370	-0.0211	-0.2256	0.0162	0.0317	-0.1962	-0.2156
3	$CH(CH_3)CH(CH_3)$	****	****	****	****	****	****	****	****	****	****	****	****	****
4	$CH(CH_3)C(CH_3)_2$	****	****	****	****	****	****	****	****	****	****	****	****	****
5	$C(CH_3)_2C(CH_3)_2$	****	****	****	****	****	****	****	****	****	****	****	****	****
6	CH_n=CH_m-CH_p=CH_k (*k,m,n,p* in 0..2)	0.2736	-0.0427	-0.3420	-0.5617	-0.5695	-0.2080	-0.2673	-0.1188	-0.1908	-0.3873	-0.2523	-0.2213	0.2736
7	CH_3-CH_m=CH_n (*m,n* in 0..2)	-0.1712	-0.0231	0.2339	0.0173	0.2614	0.0113	0.2153	-0.0860	0.2118	0.1161	0.4139	-0.0296	-0.1712
8	CH_2-CH_m=CH_n (*m,n* in 0..2)	-0.3032	0.1616	-0.0889	0.7139	-0.3033	0.3887	-0.5355	0.3290	-0.2445	0.3041	-0.3880	0.2059	-0.3032
9	CH_p-CH_m=CH_n (*m,n* in 0..2; *p* in 0..1)	0.2543	-0.5451	-0.2249	-0.5207	0.1906	-0.6964	0.0219	-0.3310	-0.0087	-0.6388	-0.2203	-0.5026	0.2543
10	CHCHO or CCHO	****	****	****	****	****	****	****	****	****	****	****	****	****
11	CH_3COCH_2	0.3730	-0.1831	0.3565	0.0548	0.4955	-0.0975	0.8558	0.0463	0.0577	-0.0729	0.2147	-0.1488	0.3730
12	CH_3COCH or CH_3COC	****	****	****	****	****	****	****	****	****	****	****	****	****
13	CHCOOH or CCOOH	0.0349	-0.2373	-0.1692	-0.3438	-0.0496	-0.1837	0.0025	-0.4877	-0.0280	-0.1855	-0.2519	-0.1631	0.0349
14	CH_3COOCH or CH_3COOC	-0.0031	****	0.1004	****	-0.6116	****	-0.2886	****	-0.1639	****	-0.0906	****	-0.0031
15	CO-O-CO	****	****	****	****	****	****	****	****	****	****	****	****	****
16	CHOH	-0.0202	-0.2409	0.0717	-0.0102	-0.2283	-0.3003	-0.0856	-0.2054	0.0897	-0.1083	0.0650	-0.1746	-0.0202
17	COH	0.2691	-0.3445	-0.3730	-0.7756	-0.1990	-0.3745	-0.1655	-1.0201	0.2900	-0.4718	0.1695	-0.1091	0.2691
18	CH_3COCH_nOH (*n* in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
19	NCCHOH or NCCOH	****	****	****	****	****	****	****	****	****	****	****	****	****
20	OH-CH_n-COO (*n* in 0..2)	-0.1302	-0.0724	-0.9562	-0.5261	0.0508	-0.3016	-1.0696	-0.2605	-0.6475	-0.0184	-1.2065	0.1851	-0.1302
21	$CH_m(OH)CH_n(OH)$ (*m,n* in 0..2)	-0.6337	****	0.4786	0.0674	-2.0125	****	-0.0382	****	-0.3093	****	-0.2551	****	-0.6337
22	$CH_m(OH)CH_n(NH_p)$ (*m,n,p* in 0..2)	-0.6256	****	-0.9371	****	-0.7020	****	-0.8702	****	-0.9641	****	-1.0751	****	-0.6256
23	$CH_m(NH_2)CH_n(NH_2)$ (*m,n* in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
24	$CH_m(NH)CH_n(NH_2)$ (*m,n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
25	$H_2NCOCH_nCH_mCONH_2$ (*m,n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
26	$CH_m(NH_n)$-COOH (*m,n* in 0..2)	0.2262	0.0000	0.3603	0.0000	0.3488	0.0000	0.0218	****	0.4185	0.0000	0.5089	0.0000	0.2262
27	HOOC-CH_n-COOH (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
28	HOOC-CH_n-CH_m-COOH (*n, m* in	****	****	****	****	****	****	****	****	****	****	****	****	****
29	HO-CH_n-COOH (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
30	NH_2-CH_n-CH_m-COOH (*n, m* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
31	CH_3-O-CH_n-COOH (n in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
32	HS-CH-COOH	****	****	****	****	****	****	****	****	****	****	****	****	****
33	HS-CH_n-CH_m-COOH (*n, m* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
34	NC-CH_n-CH_m-CN (*n, m* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****

35	OH-CH_n-CH_m-CN (*n, m* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
36	HS-CH_n-CH_m-SH (*n, m* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
37	COO-CH_n-CH_m-OOC (*n, m* in 1..2)	0.4469	****	0.6427	****	0.6875	****	0.0983	****	0.4387	****	0.4819	****	0.4469
38	OOC-CH_m-CH_m-COO (*n, m* in 1..2)	-0.6981	0.4660	-0.7761	0.6166	-0.7811	0.8178	-0.0708	0.5411	0.3322	1.2502	-0.3033	1.3833	-0.6981
39	NC-CH_n-COO (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
40	$COCH_n$COO (*n* in 1..2)	-1.8650	****	-1.7823	****	-2.4776	****	-4.2789	****	-0.2887	****	-1.0733	****	-1.8650
41	CH_m-O-CH_n=CH_p (*m,n,p* in 0..3)	****	****	****	****	****	****	****	****	****	****	****	****	****
42	CH_m=CH_n-F (*m,n* in 0..2)	-0.0771	****	-0.0802	****	-0.0851	****	-0.0019	****	-0.0470	****	-0.0282	****	-0.0771
43	CH_m=CH_n-Br (*m,n* in 0..2)	****	-0.1566	****	-0.0300	****	-0.3653	****	-0.3419	****	-0.2270	****	-0.0672	****
44	CH_m=CH_n-I (*m,n* in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
45	CH_m=CH_n-Cl (*m,n* in 0..2)	0.1340	-0.0433	0.1338	0.0066	0.3215	-0.0546	0.1440	-0.1301	0.2013	-0.0126	0.2409	-0.0067	0.1340
46	CH_m=CH_n-CN (*m,n* in 0..2)	-0.3279	-0.7947	-0.1976	-0.9862	-0.4389	-0.7243	-0.9403	-0.8521	-0.6087	-1.0236	-0.5363	-0.9052	-0.3279
47	CH_n=CH_m-COO-CH_p (*m,n,p* in 0..3)	0.0069	0.0317	-0.2782	-0.2842	-0.2485	-0.2269	0.1015	0.0910	-0.1050	-0.0816	0.0177	0.0795	0.0069
48	CH_m=CH_n-CHO (*m,n* in 0..2)	-0.6932	-1.5263	-0.0810	-1.6780	0.0021	-0.6885	0.3819	-2.2257	-0.3543	-1.2512	-0.5143	-1.1984	-0.6932
49	CH_m=CH_n-COOH (*m,n* in 0..2)	-0.8306	****	-0.8969	****	-1.0459	****	-1.2008	****	-0.9549	****	-0.8379	****	-0.8306
50	aC-CH_n-X (*n* in 1..2) X: Halogen	-0.5178	****	-0.6314	****	-0.3923	****	-0.7383	****	-0.5454	****	-0.5294	****	-0.5178
51	aC-CH_n-NH_m (*n* in 1..2; *m* in 0..2))	0.4471	****	0.2560	****	0.1222	****	-0.0642	****	0.3025	****	0.5023	****	0.4471
52	aC-CH_n-O- (*n* in 1..2)	0.6104	0.1555	0.4508	0.1158	0.3795	-0.2851	0.5760	-0.4067	0.3126	0.2177	0.4127	0.1494	0.6104
53	aC-CH_n-OH (*n* in 1..2)	0.1408	-0.0105	0.4880	-0.0293	0.3236	0.0274	0.5686	0.1099	0.3018	-0.0632	0.3113	-0.1477	0.1408
54	aC-CH_n-CN (*n* in 1..2)	****	0.0051	****	0.1590	****	0.0594	****	-0.0081	****	0.2819	****	0.0526	****
55	aC-CH_n-CHO (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
56	aC-CH_n-SH (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
57	aC-CH_n-COOH (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
58	aC-CH_n-CO- (*n* in 1..2)	****	-0.0357	****	0.0104	****	0.2390	****	0.1591	****	0.0396	****	0.0493	****
59	aC-CH_n-S- (*n* in 1..2)	****	-0.2919	****	-0.6351	****	-0.1592	****	-0.2046	****	0.3179	****	0.3396	****
60	aC-CH_n-OOC-H (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
61	aC-CH_m-NO_2 (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
62	aC-CH_n-$CONH_2$ (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
63	aC-CH_n-OOC (*n* in 1..2)	0.4686	0.0303	0.4189	0.0462	0.3224	0.0302	0.5022	0.1718	0.3665	-0.2482	0.3405	-0.0679	0.4686
64	aC-CH_n-COO (*n* in 1..2)	0.3346	-0.0228	0.5036	0.1422	0.0694	-0.1243	0.3202	-0.2149	0.4888	-0.3434	0.4867	-0.4429	0.3346
65	aC-SO_2-OH	****	****	****	****	****	****	****	****	****	****	****	****	****
66	aC-$CH(CH_3)_2$	-1.3901	-0.0910	-1.8195	-0.4011	-1.4011	0.0001	-0.4869	0.2649	-1.1552	-0.5553	-0.9422	-0.0808	-1.3901
67	aC-$C(CH_3)_3$	0.2991	-0.0127	0.2530	0.2438	0.3429	0.0812	0.2798	1.0165	0.3542	0.5184	0.4483	-0.2791	0.2991
68	aC-CF_3	0.1531	0.0424	0.3470	0.1010	0.2803	0.1118	-0.0413	0.1208	0.0848	0.0624	0.0273	-0.0074	0.1531
69	$(CH_n{=}C)_{cyc}$-CHO (*n* in 0..2)	1.0675	0.3007	0.2864	0.4844	0.3154	-0.2639	0.2947	0.6183	0.6517	-0.1476	0.6216	-0.0205	1.0675
70	$(CH_n{=}C)_{cyc}$-COO-CH_m (*n,m* in 0..3)	****	0.3055	****	0.2118	****	0.1074	****	0.1839	****	-0.4792	****	-0.4464	****
71	$(CH_n{=}C)_{cyc}$-CO- (*n* in 0..2)	****	0.5239	****	0.3025	****	0.2088	****	0.0438	****	0.5817	****	0.4229	****
72	$(CH_n{=}C)_{cyc}$-CH_3 (*n* in 0..2)	-0.0655	0.0845	0.3278	0.2269	0.4787	0.2541	-0.0549	0.6388	0.2507	0.2678	0.3649	0.2391	-0.0655
73	$(CH_n{=}C)_{cyc}$-CH_2 (*n* in 0..2)	0.4720	-0.3818	0.0573	-0.4919	0.1077	-0.5961	0.4593	-0.6096	-0.4267	-0.2819	-0.2841	-0.3409	0.4720
74	$(CH_n{=}C)_{cyc}$-CN (*n* in 0..2)	****	0.1111	****	0.0100	****	-0.1014	****	-0.2254	****	0.0394	****	0.2240	****

75	$(CH_n{=}C)_{cyc}$-Cl (*n* in 0..2)	-0.0174	0.0085	-0.0040	0.0654	-0.2488	0.0744	-0.3374	-0.0197	-0.1947	0.0863	-0.1816	0.0594	-0.0174
76	CH_{cyc}-CH_3	0.1617	-0.0418	-0.1039	-0.1943	-0.0965	-0.0070	-0.0103	-0.0280	-0.1309	0.1500	-0.0806	-0.1868	0.1617
77	CH_{cyc}-CH_2	0.0635	0.2944	0.1360	0.3422	0.0175	0.3337	-0.0708	0.5201	0.1780	0.3558	0.1664	0.2557	0.0635
78	CH_{cyc}-CH	****	-0.2875	****	0.9491	****	-0.1082	****	1.4752	****	0.5631	****	0.0253	****
79	CH_{cyc}-C	****	****	****	****	****	****	****	****	****	****	****	****	****
80	CH_{cyc}-CH=CH_n (*n* in 1..2)	1.3088	0.2967	1.7672	0.9563	1.1725	0.4304	1.1241	0.5865	0.9590	0.6805	0.5345	0.6288	1.3088
81	CH_{cyc}-C=CH_n (*n* in 1..2)	0.2406	0.5031	****	0.3233	-1.3843	0.7125	-0.4182	1.0504	****	0.4853	-2.2284	0.4768	0.2406
82	CH_{cyc}-Cl	0.0257	-0.0812	-0.0001	-0.1057	0.0151	-0.1684	0.0731	0.0380	-0.0066	-0.1851	0.0001	-0.1387	0.0257
83	CH_{cyc}-F	****	****	****	****	****	****	****	****	****	****	****	****	****
84	CH_{cyc}-OH	0.0797	****	0.2227	****	0.5691	****	0.3085	****	0.6193	****	0.4963	****	0.0797
85	CH_{cyc}-NH_2	****	-0.1407	****	0.5338	****	-0.1351	****	-0.3309	****	-0.6419	****	-0.4300	****
86	CH_{cyc}-NH-CH_n (*n* in 0..3)	****	-0.1128	****	-0.3077	****	-0.2112	****	-0.1903	****	-0.0236	****	-0.2425	****
87	CH_{cyc}-N-CH_n (*n* in 0..3)	-0.0580	****	-0.2114	****	-1.0213	****	-0.8909	****	-0.3087	****	-1.0499	****	-0.0580
88	CH_{cyc}-SH	****	****	****	****	****	****	****	****	****	****	****	****	****
89	CH_{cyc}-CN	****	****	****	****	****	****	****	****	****	****	****	****	****
90	CH_{cyc}-COOH	0.2481	****	0.3285	****	0.6706	****	1.2973	****	0.7755	****	0.6862	****	0.2481
91	CH_{cyc}-CO	0.8042	****	0.8784	****	1.6918	****	1.6952	****	1.6515	****	1.9332	****	0.8042
92	CH_{cyc}-NO_2	****	****	****	****	****	****	****	****	****	****	****	****	****
93	CH_{cyc}-S-	****	0.4048	****	0.1410	****	0.4632	****	0.3856	****	0.3079	****	0.3541	****
94	CH_{cyc}-CHO	****	****	****	****	****	****	****	****	****	****	****	****	****
95	CH_{cyc}-O-	0.6481	0.0368	0.6856	-0.1185	-0.3891	-0.0508	0.0841	-0.5279	-0.0898	-0.2893	-0.0996	0.0577	0.6481
96	CH_{cyc}-OOCH	****	****	****	****	****	****	****	****	****	****	****	****	****
97	CH_{cyc}-COO	-1.4180	0.2074	-0.9837	-0.4858	0.3256	0.5251	0.6104	0.1570	0.5905	0.2617	0.7028	0.0426	-1.4180
98	CH_{cyc}-OOC	-0.5002	****	-0.6494	****	0.0684	****	-0.5904	****	-0.6940	****	-0.6726	****	-0.5002
99	C_{cyc}-CH_3	0.4925	0.0098	0.3342	0.0934	0.2086	-0.0724	0.3966	0.0352	0.1006	-0.1272	0.1656	-0.0368	0.4925
100	C_{cyc}-CH_2	-0.1094	-0.4090	-0.1517	-0.2691	-0.1147	-0.4371	-0.0568	-0.6990	-0.2259	-0.0504	-0.2193	-0.0427	-0.1094
101	C_{cyc}-OH	-0.7038	****	-0.8063	****	-0.6692	****	-0.6998	****	-0.5360	****	-0.5070	****	-0.7038
102	$>N_{cyc}$-CH_3	-1.2066	0.1103	-0.2019	0.2099	-0.4868	0.1756	0.0853	0.6472	0.0919	0.0726	-0.2492	0.1523	-1.2066
103	$>N_{cyc}$-CH_2	-0.7739	0.3840	-0.6276	0.2840	-0.7202	0.4209	-1.8555	0.4905	-0.5019	0.3356	-0.6378	0.2710	-0.7739
104	AROMRINGs$^1s^2$	-0.1769	0.1723	-0.1816	0.2097	-0.1361	0.1386	-0.0783	0.1272	-0.1280	0.1021	-0.0969	0.1780	-0.1769
105	AROMRINGs$^1s^3$	-0.1773	-0.1755	0.1415	-0.3978	-0.0428	-0.2394	-0.2547	-0.3629	-0.4406	-0.1069	-0.1728	-0.1209	-0.1773
106	AROMRINGs$^1s^4$	-0.0844	-0.0608	-0.1789	-0.0067	-0.1036	0.0204	-0.0250	0.0392	-0.1419	0.0602	-0.1504	0.1637	-0.0844
107	AROMRINGs$^1s^2s^3$	-0.8970	-0.2962	-1.0224	-0.4284	-1.3342	-0.6484	-1.2374	-1.0101	-1.3327	-0.6523	-1.2496	-0.4397	-0.8970
108	AROMRINGs$^1s^2s^4$	0.0620	0.0797	0.0487	-0.1050	0.0410	-0.1312	0.1380	-0.1701	-0.0209	-0.1722	-0.0483	-0.1165	0.0620
109	AROMRINGs$^1s^3s^5$	0.9750	0.1039	0.5147	-0.4384	0.3431	0.0309	0.4799	-0.4711	0.9596	-0.0603	0.9389	-0.0419	0.9750
110	AROMRINGs$^1s^2s^3s^4$	-0.3576	-0.1394	-0.5347	-0.0329	-0.6996	-0.2504	-0.2893	-1.2009	-0.4443	0.0538	-0.3641	-0.0909	-0.3576
111	AROMRINGs$^1s^2s^3s^5$	0.3356	-0.0381	0.3463	0.0338	0.2977	-0.0871	0.3619	-0.2276	0.5856	-0.0414	0.4694	-0.0448	0.3356
112	AROMRINGs$^1s^2s^4s^5$	0.4565	0.0981	0.2489	0.2706	0.0031	0.0371	0.1238	0.5343	-0.1304	-0.1554	-0.0715	0.1909	0.4565
113	PYRIDINEs^2	-3.5683	****	-1.8976	****	-1.7913	****	-3.4187	****	-1.3125	****	-2.0497	****	-3.5683
114	PYRIDINEs^3	0.6433	****	0.4661	****	0.3197	****	0.6967	****	0.5743	****	0.4823	****	0.6433

115	PYRIDINEs^{4}	0.4128	****	0.3197	****	-0.1549	****	0.2061	****	0.1359	****	-0.1956	****	0.4128
116	PYRIDINEs^{2}s^{3}	****	****	****	****	****	****	****	****	****	****	****	****	****
117	PYRIDINEs^{2}s^{4}	0.5067	****	0.6051	****	0.5708	****	0.9500	****	0.4179	****	0.4536	****	0.5067
118	PYRIDINEs^{2}s^{5}	****	-0.2465	****	-0.2867	****	-0.0442	****	0.5791	****	0.0300	****	0.2084	****
119	PYRIDINEs^{2}s^{6}	****	****	****	****	****	****	****	****	****	****	****	****	****
120	PYRIDINEs^{3}s^{4}	****	****	****	****	****	****	****	****	****	****	****	****	****
121	PYRIDINEs^{3}s^{5}	****	****	****	****	****	****	****	****	****	****	****	****	****
122	PYRIDINEs^{2}s^{3}s^{6}	****	****	****	****	****	****	****	****	****	****	****	****	****
123	(CH^{n}=CH^{m})^{cyc}-COOH	1.9335	****	1.3822	****	1.4197	****	3.0471	****	0.5406	****	1.0354	****	1.9335
124	AROMRINGs^{1}s^{2}s^{3}s^{4}s^{5}	-1.4672	-0.2969	-1.5435	0.0898	-1.1716	0.4501	-1.2833	0.7438	-3.0130	0.3279	-2.1132	-0.1988	-1.4672
125	aC-NHCOCH$_2$N	****	****	****	****	****	****	****	****	****	****	****	****	****
126	(N=C)$_{cyc}$-CH$_3$	1.4459	0.0987	0.6505	0.2277	0.8497	0.2606	0.8816	-0.3916	0.3579	0.4891	0.5249	0.5919	1.4459
127	aC-CONH(CH$_2$)$_2$N	****	****	****	****	****	****	****	****	****	****	****	****	****
128	aC-SO$_2$NH$_n$ (n>=0;n<3)	1.0075	-0.0458	0.8867	-0.0755	0.7118	0.0687	0.5428	0.1131	0.7967	0.0251	0.9353	-0.0841	1.0075
129	aC-SO$_2$NH$_n$ (n>=0;n<3)	0.7076	****	-0.3532	****	-0.1085	****	-0.3788	****	-0.6173	****	-0.6392	****	0.7076
130	aC-SO$_2$NH$_n$ (n>=0;n<3)	0.5896	****	0.7720	****	0.8423	****	0.5188	****	1.1647	****	1.2373	****	0.5896

[a] The symbols $EUA_{C\ 2j}$, $EUA_{NC\ 2j}$, $ERA_{C\ 2j}$, $ERA_{NC\ 2j}$, $EFW_{C\ 2j}$, $EFW_{NC\ 2j}$, $ESW_{C\ 2j}$, $ESW_{NC\ 2j}$, $ENS_{C\ 2j}$, $ENS_{NC\ 2j}$, $EAS_{C\ 2j}$, and $EAS_{NC\ 2j}$ represent the contributions (D_j) of the second-order groups for the corresponding properties.

Table B4. MG Method Based Property Models Analysed Using Step-Wise Regression Method: Third-Order Groups and their contributions [a] for the Properties÷ LC_{50}(FM), LC_{50}(DM), LD_{50}, $LogW_s$, BCF, PEL(OSHA-TWA), and PCO

	Group	$LC50(FM)_{3k}$	$LC50(DM)_{3k}$	$LD50_{3k}$	$LogWs_{3k}$	BCF_{3k}	PEL_{3k}	PCO_{3k}
1	HOOC-$(CH_n)_m$-COOH (m>2, n in 0..2)	0.0498	****	-0.4908	0.1790	****	-0.4810	-0.0249
2	NH_n-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	****	1.0484	****	****	****
3	NH_2-$(CH_n)_m$-OH (m>2, n in 0..2)	****	****	-0.2787	****	****	****	****
4	OH-$(CH_n)_m$-OH (m>2, n in 0..2)	****	****	0.1759	1.3600	****	****	-0.1240
5	OH-$(CH_p)_k$-O-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****
6	OH-$(CH_p)_k$-S-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****
7	OH-$(CH_p)_k$-NH_x-$(CH_n)_m$-OH (m,k>0; p,n,x in 0..2)	****	****	****	****	****	****	****
8	CH_p-O-$(CH_n)_m$-OH (m>2; n,p in 0..2)	****	****	****	****	****	****	****
9	NH_2-$(CH_n)_m$-NH_2 (m>2; n in 0..2))	-0.6399	****	0.3552	****	****	0.5413	****
10	NH_k-$(CH_n)_m$-NH_2 (m>2; k in 0..1; n in 0..2)	****	****	****	****	****	****	****
11	SH-$(CH_n)_m$-SH (m>2; n in 0..2)	****	****	****	****	****	****	****
12	NC-$(CH_n)_m$-CN (m>2)	-0.7496	****	1.1305	1.0800	****	-0.1657	****
13	COO-$(CH_n)_m$-OOC (m>2; n in 0..2)	****	****	****	****	****	****	****
14	aC-$(CH_n$=$CH_m)_{cyc}$ (fused rings) (n,m in 0..1)	-0.1522	-0.3984	0.1430	0.6020	-0.1148	-0.8684	0.5575
15	aC-aC (different rings)	0.1281	0.2825	0.0735	-0.2230	0.0330	0.0122	****
16	aC-CH_{ncyc} (different rings) (n in 0..1)	-0.1468	0.0000	0.1622	0.5020	-0.1038	****	****
17	aC-CH_{ncyc} (fused rings) (n in 0..1)	-0.2569	-0.0249	0.0701	0.3690	-0.0008	0.5251	-0.0309
18	aC-$(CH_n)_m$-aC (different rings) (m>1; n in 0..2)	0.0710	****	-0.4276	-1.3000	-2.5311	****	****
19	aC-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	0.9803	0.0000	-0.0149	-0.3336	****	****	****
20	CH_{cyc}-CH_{cyc} (different rings)	****	****	-0.2107	0.6122	****	****	0.2449
21	CH_{cyc}-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****
22	$CH_{multiring}$	0.0499	-0.0154	-0.0344	-0.0247	0.2805	0.1166	0.0582
23	$C_{multiring}$	-0.0662	-0.0052	0.0023	0.0710	-0.3160	-0.2036	-0.0162
24	aC-CH_m-aC (different rings) (m in 0..2)	-0.2276	-0.4233	-0.0873	0.2310	0.0506	-0.9455	0.0509
25	aC-(CH_m=CH_n)-aC (different rings) (m,n in 0..2)	****	-0.5224	-0.3410	-0.8542	-0.0416	****	****
26	$(CH_m$=C$)_{cyc}$-CH=CH-$(C$=$CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****
27	$(CH_m$=C$)_{cyc}$-CH_p-$(C$=$CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****
28	aC-CO-aC (different rings)	-0.2159	****	0.1961	1.0600	-0.4738	****	****
29	aC-CH_m-CO-aC (different rings) (m in 0..2)	****	****	-0.3475	-0.5809	0.5000	****	****
30	aC-CO-$(C$=$CH_n)_{cyc}$ (different rings) (n in 0..1)	****	****	-0.0519	-0.4667	****	****	****
31	aC-CO-CO-aC (different rings)	****	****	****	0.1407	****	****	****

32	aC-CO_{cyc} (fused rings)	0.2092	0.2980	-0.0138	0.0380	-0.0285	-1.1411	-0.1487
33	aC-CO-$(CH_n)_m$-CO-aC (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****
34	aC-CO-CH_{ncyc} (different rings) (n in 0..1)	****	****	****	-0.6102	****	****	****
35	aC-CO-NH_n-aC (different rings) (n in 0..1)	-0.0607	****	-0.2718	-0.5870	-0.2114	****	****
36	aC-NH_nCONH_m-aC (different rings) (n,m in 0..1)	1.4228	-0.0895	****	0.8040	-0.6021	****	****
37	aC-CO-N_{cyc} (different rings)	****	****	0.3225	-1.0070	****	****	****
38	aC-S_{cyc} (fused rings)	-0.8909	-0.1716	0.0828	0.1650	-0.0918	-0.5763	****
39	aC-S-aC (different rings)	-1.9636	****	-0.5044	0.4514	-1.7418	****	****
40	aC-PO_n-aC (different rings) (n in 0..4)	****	****	****	****	****	****	****
41	aC-SO_n-aC (different rings) (n in 1..4)	****	0.1267	0.0683	-0.1410	****	****	****
42	aC-NH_{ncyc} (fused rings) (n in 0..1)	0.2702	0.9450	-0.0298	-0.1200	0.0486	0.5000	****
43	aC-NH-aC (different rings)	0.2900	****	-0.0049	-0.1120	0.2117	-0.5446	****
44	aC-$(C=N)_{cyc}$ (different rings)	****	****	-0.1266	0.5247	-0.3031	****	****
45	aC-$(N=CH_n)_{cyc}$ (fused rings) (n in 0..1)	-0.4338	-0.5251	0.0640	-0.0249	-0.1578	****	****
46	aC-$(CH_n=N)_{cyc}$ (fused rings) (n in 0..1)	****	****	-0.1276	0.4628	****	****	****
47	aC-O-CH_n-aC (different rings) (n in 0..2)	****	****	-0.2231	0.8493	****	****	****
48	aC-O-aC (different rings)	0.1437	0.0020	-0.0688	-0.3224	-0.0149	-0.1030	****
49	aC-CH_n-O-CH_m-aC (different rings) (n,m in 0..2)	****	****	-0.7009	0.2160	****	****	****
50	aC-O_{cyc} (fused rings)	-0.4414	-0.5119	0.0329	-0.3410	-0.0203	****	****
51	AROM.FUSED[2]	0.0032	0.1720	-0.1119	-0.0962	0.0643	0.0763	-0.0183
52	AROM.FUSED[2]s^1	0.1792	0.4007	0.2156	0.1030	0.2267	-0.1548	-0.0209
53	AROM.FUSED[2]s^2	0.2418	-0.1498	-0.0803	-0.0515	-0.0089	1.8721	0.0295
54	AROM.FUSED[2]s^2s^3	-0.2860	1.2197	0.3000	0.1132	0.2159	****	0.1743
55	AROM.FUSED[2]s^1s^4	-0.2961	****	-0.2363	-0.2116	-0.0322	****	****
56	AROM.FUSED[2]s^1s^2	****	****	0.1553	0.0217	-0.4516	****	****
57	AROM.FUSED[2]s^1s^3	****	****	0.3250	-0.0558	0.2075	****	****
58	AROM.FUSED[3]	0.0955	-0.3703	-0.1814	-0.2256	0.0443	0.1575	****
59	AROM.FUSED[4a]	****	****	-0.2323	-0.7000	-0.0350	-0.1471	****
60	AROM.FUSED[4a]s^1	****	****	-0.4303	0.3143	-0.0103	****	****
61	AROM.FUSED[4a]s^1s^4	****	****	1.1705	0.3697	****	****	****
62	AROM.FUSED[4p]	1.5851	-0.0972	-0.1580	0.0143	-0.1692	-0.0713	****
63	AROM.FUSED[4p]s^3s^4	****	****	0.2203	1.8369	****	****	****
64	PYRIDINE.FUSED[2]	-0.0274	-0.4928	0.0166	0.4989	-1.4055	****	****
65	PYRIDINE.FUSED[2-iso]	****	-0.3098	0.6246	1.3160	****	****	****
66	PYRIDINE.FUSED[4]	0.7501	-0.0203	0.1412	1.9800	-0.3701	-0.3792	****
67	aC-N-CH_{cyc} (different rings)	****	****	0.1538	-0.1524	****	****	****
68	N multiring	-0.0136	0.6590	-0.0030	0.1349	****	****	0.0258

69	N_{cyc}-$(CH_2)_3$-N_{cyc} (different rings)	****	****	-0.0732	-0.1746	****	****	****
70	aC-$COCH_2CH_2$-aC (different rings)	****	****	-0.3699	-0.6448	****	****	****
71	aC-O-$(CH_2)_2$-N_{cyc} (different rings)	****	****	-0.5047	****	****	****	****
72	aC-CH(OH)$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****
73	Ncyc-$(CH_2)_2$-CH_{cyc} (different rings)	****	****	-0.1872	****	****	****	****
74	aC-$CONHCH_2$-CH_{cyc} (different rings)	****	****	0.5619	****	****	****	****

[a] The symbols $LC50(FM)_{3k}$, $LC50(DM)_{3k}$, $LD50_{3k}$, $LogWs_{3k}$, $BCF_{3k,}$ $PEL_{3k,}$ PCO_{3k} represent the contributions (E_k) of the third-order groups for the corresponding properties. Note that there are no third-order contributions for GWP, ODP and AP since these properties belong to chemicals whose structure is defined by first-order groups alone.

Table B4 (continued). MG Method Based Property Models Analysed Using Step-Wise Regression Method: Third-Order Groups and their Contributions [a] for the Properties÷ EUA_C, EUA_{NC}, ERA_C, ERA_{NC}, EFW_C, EFW_{NC}, ESW_C, ESW_{NC}, ENS_C, ENS_{NC}, EAS_C, and EAS_{NC}.

	Group	$EUA_{C\ 3k}$	EUA_{NC3k}	$ERA_{C\ 3k}$	ERA_{NC3k}	$EFW_{C\ 3k}$	$EFW_{NC\ 3k}$	$ESW_{C\ 2j}$	$ESW_{NC\ 3k}$	$ENS_{C\ 3k}$	$ENS_{NC\ 3k}$	$EAS_{C\ 3k}$	$EAS_{NC\ 3k}$	$EUA_{C\ 3k}$
1	HOOC-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
2	NH_n-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
3	NH_2-$(CH_n)_m$-OH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
4	OH-$(CH_n)_m$-OH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
5	OH-$(CH_p)_k$-O-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
6	OH-$(CH_p)_k$-S-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
7	OH-$(CH_p)_k$.NH_x-$(CH_n)_m$-OH (m,k>0; p,n,x in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
8	CH_p-O-$(CH_n)_m$-OH (m>2; n,p in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
9	NH_2-$(CH_n)_m$-NH_2 (m>2; n in 0..2))	****	****	****	****	****	****	****	****	****	****	****	****	****
10	NH_k-$(CH_n)_m$-NH_2 (m>2; k in 0..1; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
11	SH-$(CH_n)_m$.SH (m>2; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
12	NC-$(CH_n)_m$-CN (m>2)	****	****	****	****	****	****	****	****	****	****	****	****	****
13	COO-$(CH_n)_m$-OOC (m>2; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
14	aC-$(CH_n{=}CH_m)_{cyc}$ (fused rings) (n,m in 0..1)	-0.1215	****	-0.1473	****	0.1250	****	-0.0685	****	0.2146	****	0.2365	****	-0.1215
15	aC-aC (different rings)	-0.2468	0.0671	-0.1842	-0.2102	-0.2456	-0.0737	-0.3819	-0.1675	-0.1974	-0.0488	-0.0378	-0.0141	-0.2468
16	aC-CH_{ncyc} (different rings) (n in 0..1)	0.2361	0.3257	0.1430	0.3934	0.3252	0.4016	0.2774	0.7521	-0.0755	0.1389	-0.1467	0.0749	0.2361
17	aC-CH_{ncyc} (fused rings) (n in 0..1)	0.1459	0.2401	-0.0771	0.5147	0.0586	0.3469	0.2672	0.2325	-0.1877	0.2791	-0.1218	0.3119	0.1459
18	aC-$(CH_n)_m$-aC (different rings) (m>1; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
19	aC-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
20	CH_{cyc}-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
21	CH_{cyc}-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
22	$CH_{multiring}$	0.0300	0.0278	0.0391	-0.0003	-0.0363	-0.0292	-0.0287	-0.0117	-0.0504	-0.0425	-0.0394	-0.0300	0.0300
23	$C_{multiring}$	-0.0165	-0.0191	-0.0075	-0.0090	0.0119	0.0001	0.0147	-0.0020	0.0377	0.0147	0.0267	0.0038	-0.0165
24	aC-CH_m-aC (different rings) (m in 0..2)	-0.3033	0.0545	-0.2308	0.0585	-0.1917	0.0840	-0.3588	0.0777	0.1200	0.1306	0.0679	0.1377	-0.3033
25	aC-$(CH_m{=}CH_n)$-aC (different rings) (m,n in 0..2)	-1.2865	****	-1.5110	****	-0.9390	****	-1.3848	****	-0.2259	****	-0.3749	****	-1.2865
26	$(CH_m{=}C)_{cyc}$-CH=CH-$(C{=}CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
27	$(CH_m{=}C)_{cyc}$-CH_p-$(C{=}CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
28	aC-CO-aC (different rings)	0.3796	****	0.7024	****	0.7866	****	0.5009	****	0.7473	****	0.7154	****	0.3796
29	aC-CH_m-CO-aC (different rings) (m in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
30	aC-CO-$(C{=}CH_n)_{cyc}$ (different rings) (n in 0..1)	****	****	****	****	****	****	****	****	****	****	****	****	****
31	aC-CO-CO-aC (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****

32	aC-CO_{cyc} (fused rings)	-0.0933	0.0876	-0.0655	0.0969	-0.0074	0.1370	-0.1558	0.2045	-0.0724	-0.1393	-0.0848	-0.1407	-0.0933
33	aC-CO-$(CH_n)_m$-CO-aC (different rings) (m>0; n in	****	****	****	****	****	****	****	****	****	****	****	****	****
34	aC-CO-CH_{ncyc} (different rings) (n in 0..1)	****	****	****	****	****	****	****	****	****	****	****	****	****
35	aC-CO-NH_n-aC (different rings) (n in 0..1)	****	-0.1635	****	0.0182	****	-0.1253	****	-0.0293	****	-0.2030	****	-0.2010	****
36	aC-NH_nCONH_m-aC (different rings) (n,m in 0..1)	0.0224	****	0.1302	****	0.0626	****	-0.1130	****	0.1627	****	0.1987	****	0.0224
37	aC-CO-N_{cyc}(different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
38	aC-S_{cyc} (fused rings)	1.0646	0.5747	0.9674	1.0012	1.0749	0.8358	1.1511	0.4969	1.7345	0.1996	1.7415	0.4978	1.0646
39	aC-S-aC (different rings)	0.1689	****	0.3578	****	0.2071	****	0.0500	****	0.2837	****	0.3008	****	0.1689
40	aC-PO_n-aC (different rings) (n in 0..4)	****	****	****	****	****	****	****	****	****	****	****	****	****
41	aC-SO_n-aC (different rings) (n in 1..4)	-0.6611	****	-0.3872	****	-0.2601	****	-0.1723	****	-0.2444	****	-0.4162	****	-0.6611
42	aC-NH_{ncyc}(fused rings) (n in 0..1)	0.2286	0.5134	0.0775	0.6853	0.1433	0.5197	0.3180	0.2492	0.3238	0.4479	0.2342	0.6158	0.2286
43	aC-NH-aC (different rings)	3.0372	0.4637	2.9764	0.9832	2.6222	0.2142	2.1095	0.2696	2.7532	0.3378	2.6061	0.2319	3.0372
44	aC-$(C{=}N)_{cyc}$ (different rings)	0.3105	0.5157	0.0364	0.3922	-0.3900	0.4152	0.3847	0.0866	-1.1713	-0.3841	-0.9995	-0.0549	0.3105
45	aC-$(N{=}CH_n)_{cyc}$ (fused rings) (n in 0..1)	-0.3841	-0.2992	-0.3342	-0.2709	-0.2907	-0.3074	0.1324	0.0130	-0.5998	-0.3192	-0.4553	-0.5872	-0.3841
46	aC-$(CH_n{=}N)_{cyc}$ (fused rings) (n in 0..1)	****	****	****	****	****	****	****	****	****	****	****	****	****
47	aC-O-CH_n-aC (different rings) (n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
48	aC-O-aC (different rings)	0.6499	-0.0274	0.4143	-0.0110	0.4380	-0.0296	0.5488	-0.0890	1.1196	0.0980	0.9371	0.0870	0.6499
49	aC-CH_n-O-CH_m-aC (different rings) (n,m in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
50	aC-O_{cyc} (fused rings)	0.0383	-0.6411	0.0370	-0.8716	-0.1231	-0.8194	0.0243	-0.9153	-0.0054	-0.4711	-0.0885	-0.4723	0.0383
51	AROM.FUSED[2]	0.1136	-0.2310	0.2166	-0.4164	0.0549	-0.3894	0.1091	-0.3073	0.0170	-0.1885	0.0079	-0.1675	0.1136
52	AROM.FUSED[2]s^1	0.1925	0.5089	0.0158	0.6586	-0.1485	0.8062	0.0589	0.6587	0.2571	0.3565	0.3012	0.5469	0.1925
53	AROM.FUSED[2]s^2	-0.2109	0.3013	-0.1850	0.6488	-0.0357	0.3817	-0.2511	0.4418	0.3028	0.1445	0.2705	-0.0524	-0.2109
54	AROM.FUSED[2]s^2s^3	-1.5896	1.2673	0.6844	1.8541	0.7961	1.6729	-1.6291	2.2637	-1.0883	0.7740	-0.9250	0.6817	-1.5896
55	AROM.FUSED[2]s^1s^4	-0.0973	****	0.0836	****	-0.2053	****	-0.0020	****	-0.0769	****	-0.0043	****	-0.0973
56	AROM.FUSED[2]s^1s^2	0.7666	-1.2625	-0.2823	-0.6895	-0.3718	-0.0288	0.8986	-1.2125	0.9963	2.6709	0.9649	1.8788	0.7666
57	AROM.FUSED[2]s^1s^3	-0.6426	****	-0.8108	****	-0.6896	****	-0.8285	****	-1.0441	****	-0.9791	****	-0.6426
58	AROM.FUSED[3]	-0.1308	0.1254	0.3385	0.1587	0.1683	-0.0901	-0.3007	-0.0059	0.2048	-0.1068	0.3250	-0.1080	-0.1308
59	AROM.FUSED[4a]	0.3064	****	0.7653	1.1669	0.5645	1.1140	0.5066	1.0208	0.1222	0.5681	0.6221	0.5363	0.3064
60	AROM.FUSED[4a]s^1	2.1227	****	1.8670	****	1.8023	****	2.4139	****	1.6935	****	1.9882	****	2.1227
61	AROM.FUSED[4a]s^1s^4	-1.0810	****	-2.0746	****	-1.9895	****	-2.5320	****	-1.6361	****	-1.3919	****	-1.0810
62	AROM.FUSED[4p]	-0.0590	-0.1254	-0.2886	-0.4931	-0.0769	-0.0628	0.0270	-0.1591	-0.0097	0.0760	-0.1541	0.0488	-0.0590
63	AROM.FUSED[4p]s^3s^4	0.2168	****	0.1719	****	-0.0581	****	0.5679	****	-0.1928	****	-0.0284	****	0.2168
64	PYRIDINE.FUSED[2]	-0.7523	****	-0.5874	****	-0.5573	****	0.2154	****	-0.4242	****	-0.2953	****	-0.7523
65	PYRIDINE.FUSED[2-iso]	****	****	****	****	****	****	****	****	****	****	****	****	****
66	PYRIDINE.FUSED[4]	****	****	****	****	****	****	****	****	****	****	****	****	****
67	aC-N-CH_{cyc} (different rings)	****	****	0.0000	****	0.0000	****	****	****	0.0000	****	0.0000	****	****
68	N multiring	-0.2136	-0.4632	-0.0115	-0.4748	-0.0907	-0.0129	-0.2649	0.0949	0.1129	-0.0060	0.1290	0.0151	-0.2136

69	N_{cyc}-$(CH_2)_3$-N_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
70	aC-$COCH_2CH_2$-aC (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
71	aC-O-$(CH_2)_2$-N_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
72	aC-CH(OH)$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
73	Ncyc-$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
74	aC-$CONHCH_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****

[a] The symbols $EUA_{C\ 3k}$, $EUA_{NC\ 3k}$, $ERA_{C\ 3k}$, $ERA_{NC\ 3k}$, $EFW_{C\ 3k}$, $EFW_{NC\ 3k}$, $ESW_{C\ 3k}$, $ESW_{NC\ 3k}$, $ENS_{C\ 3k}$, $ENS_{NC\ 3k}$, $EAS_{C\ 3k}$, and $EAS_{NC\ 3k}$ represent the contributions (E_k) of the third-order groups for the corresponding properties.

Table B5. MG Method Based Property Models Analysed Using Simultaneous Regression Method: First-Order Groups and their Contributions [a] for the Properties÷ LC_{50}(FM), LC_{50}(DM), LD_{50}, $LogW_{s}$,, BCF, PEL(OSHA-TWA), PCO, GWP, ODP, and AP.

	Group	$LC50(FM)_{1i}$	$LC50(DM)_{1i}$	$LD50_{1i}$	$LogWs_{1i}$	BCF_{1i}	PEL_{1i}	PCO_{1i}	GWP_{1i}	ODP_{1i}	AP_{1i}
1	CH_3	0.0556	-0.3263	-0.0620	-5.1279	0.6703	0.7059	0.1421	0.3880	-0.9453	-0.1290
2	CH_2	0.3058	0.1806	0.0206	-5.0803	0.0883	0.0742	0.0452	-1.0699	****	****
3	CH	0.2640	-0.3385	0.0834	-4.7909	-0.6251	-0.5252	-0.1351	****	****	****
4	C	-0.5591	-2.3767	0.1524	-4.9670	-1.2271	-1.1333	-0.2452	****	****	****
5	CH_2=CH	1.1915	-4.7284	0.1567	-9.5872	0.6777	2.4164	-0.3641	****	****	****
6	CH=CH	0.3986	-1.6245	0.1181	-9.5886	****	1.3391	-0.7034	****	****	****
7	CH_2=C	0.7062	-0.9381	0.1993	-9.5121	0.3400	1.1805	-0.5276	****	****	****
8	CH=C	0.9899	0.0393	0.3274	-9.4355	-0.3654	0.8654	-0.8041	****	****	****
9	C=C	1.2892	-0.9820	0.2934	-9.5320	0.5447	1.2985	-0.5092	****	****	****
10	CH_2=C=CH	****	****	****	****	****	****	****	****	****	****
11	CH_2=C=C	****	****	****	****	****	****	****	****	****	****
12	C=C=C	****	****	****	****	****	****	****	****	****	****
13	CH≡C	0.9001	****	0.0737	-8.7940	****	3.0437	-0.2592	****	****	****
14	C≡C	1.5163	3.6012	0.3346	-8.7290	****	****	-0.9453	****	****	****
15	aCH	0.1574	-0.0680	0.0200	-4.5565	0.2711	0.3182	0.0263	****	****	****
16	aC fused with aromatic ring	-0.4620	-0.7792	0.0729	-4.7557	0.0144	0.2181	-0.0470	****	****	****
17	aC fused with non-aromatic ring	0.0006	-0.7416	0.0358	-4.4965	-0.0080	0.0889	-0.0958	****	****	****
18	aC except as above	0.0893	-0.4944	0.0331	-4.4476	-0.2361	0.3662	0.8361	****	****	****
19	aN in aromatic ring	0.1432	-0.3506	0.0384	-4.4675	0.1866	2.1309	****	****	****	****
20	aC-CH_3	0.3434	0.3598	0.0616	-9.5780	0.3631	0.8824	-0.0776	****	****	****
21	aC-CH_2	0.7309	0.1167	0.1132	-9.5168	-0.0324	-0.1488	-0.1815	****	****	****
22	aC-CH	0.6725	2.8366	0.5066	-9.3337	-0.9742	0.8149	-0.2730	****	****	****
23	aC-C	0.1056	1.9398	0.5653	-9.8828	-0.8256	-5.0862	-0.4485	****	****	****
24	aC-CH=CH_2	0.7720	0.1539	0.0850	-14.1770	1.2395	0.9606	0.5094	****	****	****
25	aC-CH=CH	-0.3767	****	0.1680	-14.3619	0.2045	****	1.0404	****	****	****
26	aC-C=CH_2	****	****	-0.7064	-14.4125	-0.2210	0.2660	0.1505	****	****	****
27	aC-C≡CH	****	****	****	-13.6327	****	****	****	****	****	****
28	aC-C≡C	****	****	-0.4446	-14.2762	****	****	****	****	****	****
29	OH	-0.4918	-1.9704	-0.2449	-4.9955	0.1957	1.3297	0.0674	****	****	****
30	aC-OH	0.2408	0.1957	0.0580	-9.6506	-0.4151	1.5157	0.4048	****	-1.5842	-0.0769
31	COOH	-0.6851	-0.7502	0.0011	-14.6373	-1.3264	2.5364	0.1482	****	****	****
32	aC-COOH	-0.0709	-1.9953	0.0388	-19.3390	-1.0929	1.4314	****	****	****	****
33	CH_3CO	-0.0161	-2.5920	-0.0697	-14.3754	0.2364	0.8681	0.5458	****	****	****
34	CH_2CO	-0.5132	****	0.1381	-14.4832	-1.0508	0.9384	-0.0817	****	****	****
35	CHCO	****	****	0.4061	-14.9310	****	****	-0.0316	****	****	****
36	CCO	****	****	1.1095	-14.3234	****	-1.7554	****	****	****	****
37	aC-CO	0.4173	1.4830	0.1142	-14.0458	-0.9617	0.9302	****	****	****	****
38	CHO	0.3890	-2.3399	-0.1130	-9.0480	-0.4066	2.3638	-0.2229	****	****	****
39	aC-CHO	0.7570	0.7430	-0.0527	-14.3857	-1.2851	****	****	****	****	****
40	CH_3COO	0.4260	-1.3681	-0.1990	-19.6379	-0.0072	1.1646	0.4531	****	****	****

41	CH_2COO	0.2046	0.7330	-0.0418	-19.6496	0.0120	1.7849	0.1935	****	****	****
42	CHCOO	-0.4731	****	0.1131	-19.3341	****	****	0.1831	****	****	****
43	CCOO	-1.9971	****	0.2846	-19.4530	****	****	****	****	****	****
44	HCOO	****	****	-0.2234	-15.3161	****	1.6436	0.8105	****	****	****
45	aC-COO	0.3771	0.3883	-0.0421	-19.3085	-1.1622	0.9798	0.0416	****	****	****
46	aC-OOCH	****	****	-0.2859	****	****	****	****	****	****	****
47	aC-OOC	1.1756	****	0.3537	-19.5422	-3.7123	1.6104	****	****	****	****
48	COO except as above	0.6212	6.6088	0.1822	-14.5866	-1.2944	0.1698	-0.2389	****	****	****
49	CH_3O	-0.3622	1.1271	-0.0219	-10.0648	0.2283	2.1170	0.1751	0.1245	****	****
50	CH_2O	-0.2490	0.1565	0.1148	-9.8886	-0.2561	0.9385	-0.1353	****	****	****
51	CH-O	-0.5300	****	0.4559	-9.9280	-0.1236	-0.6058	-0.2934	-1.8521	****	****
52	C-O	****	****	0.0697	-9.7047	-0.8483	****	****	****	****	****
53	aC-O	0.0233	0.7468	0.1589	-9.6786	-0.2179	0.8384	-0.5301	****	****	****
54	CH_2NH_2	0.0884	1.8188	0.0659	-9.3586	0.3801	2.6815	-0.3332	****	****	****
55	$CHNH_2$	0.0050	0.9501	0.4614	-9.7816	****	2.2834	-0.6060	****	****	****
56	CNH_2	1.0340	****	0.5006	-8.8835	****	****	****	****	****	****
57	CH_3NH	-0.6909	0.1063	-0.0037	-9.4825	****	2.8993	-0.0143	****	****	****
58	CH_2NH	0.2411	0.1899	0.3476	-9.3334	-0.5586	1.5850	0.0159	****	****	****
59	CHNH	****	-0.1355	0.2110	-8.8354	****	1.4112	****	****	****	****
60	CH_3N	0.0656	-0.2190	0.3385	-9.5616	-0.3696	1.2989	-0.4572	****	****	****
61	CH_2N	-0.9455	-0.6699	0.4260	-9.3848	-1.4876	0.4199	-0.4197	****	****	****
62	aC-NH_2	-0.0787	1.0092	0.1302	-9.8285	-0.4375	2.4018	****	****	****	****
63	aC-NH	0.0385	1.0054	0.2339	-9.6863	-0.6826	1.4233	****	****	****	****
64	aC-N	0.0719	1.2340	0.3112	-9.7588	-0.7716	0.1870	****	****	****	****
65	NH_2 except as above	0.2122	0.3256	0.0820	-5.0409	-0.2637	2.0136	-0.4508	****	****	****
66	CH=N	2.7462	0.7708	0.1311	-9.7606	0.3204	****	****	****	****	****
67	C=N	1.5451	1.5013	0.2109	-9.1601	-1.3495	****	****	****	****	****
68	CH_2CN	0.2409	****	-0.0996	-13.5065	-0.0541	2.5521	****	****	****	****
69	CHCN	2.0033	****	0.9700	-13.6088	-0.5524	2.0880	****	****	****	****
70	CCN	1.0562	****	0.7209	-12.9959	****	0.9579	****	****	****	****
71	aC-CN	0.1265	0.3584	0.0653	-13.5175	-0.3325	****	****	****	****	****
72	CN except as above	-0.3055	0.6079	0.3035	-8.9767	1.8953	2.0420	0.5544	****	****	****
73	CH_2NCO	****	****	0.0478	****	****	3.1925	****	****	****	****
74	CHNCO	****	****	****	****	****	****	****	****	****	****
75	CNCO	****	****	****	****	****	****	****	****	****	****
76	aC-NCO	****	****	-0.1101	****	****	2.6914	0.0219	****	****	****
77	$CH2NO_2$	****	****	0.0869	-20.1394	****	1.9469	1.0619	****	****	****
78	$CHNO_2$	****	****	0.6395	-20.4173	-1.3892	1.5866	****	****	****	****
79	CNO_2	1.1653	****	1.1733	-18.9795	****	-1.0086	****	****	****	****
80	aC-NO_2	0.6761	0.1244	0.2759	-20.2293	-0.0575	2.0505	1.4104	****	****	****
81	NO_2 except as above	1.0120	0.2969	-0.0791	-16.2802	****	1.7993	****	****	****	-0.0775
82	ONO	****	****	0.8755	****	****	****	****	****	****	****
83	ONO_2	-0.0709	-0.0915	0.3217	-21.3506	****	1.9774	****	****	****	-0.0775
84	$HCON(CH_2)_2$	-0.2735	****	****	-22.4833	****	****	****	****	****	****
85	$HCONHCH_2$	****	****	****	-19.0479	****	****	****	****	****	****
86	$CONH_2$	-0.2099	3.7853	0.0938	-14.4869	****	2.9582	****	****	****	****

87	$CONHCH_3$	1.1894	0.3027	0.9970	-19.3067	-0.9446	****	****	****	****	****
88	$CONHCH_2$	1.0479	-0.2683	0.1235	-19.2945	-1.7323	****	****	****	****	****
89	$CON(CH_3)_2$	-0.7894	****	0.6593	-23.8705	2.1722	2.6902	****	****	****	****
90	$CONCH_3CH_2$	****	****	****	-23.7820	****	****	****	****	****	****
91	$CON(CH_2)_2$	-0.5577	-0.3976	0.3369	-23.5732	-1.5293	****	****	****	****	****
92	CONHCO	****	****	0.6221	-24.3563	****	****	****	****	****	****
93	CONCO	1.4698	****	0.4604	-23.7997	****	****	****	****	****	****
94	$aC\text{-}CONH_2$	-0.5257	-1.5913	0.2189	-19.1698	****	****	****	****	****	****
95	aC-NH(CO)H	0.0007	****	0.2914	-18.3036	****	****	****	****	****	****
96	aC-N(CO)H	-0.0840	****	0.2079	****	****	****	****	****	****	****
97	aC-CONH	1.3217	2.6195	0.1560	-18.7729	-1.5536	****	****	****	****	****
98	aC-NHCO	-0.2182	-0.1832	0.0072	-18.9641	-1.2849	****	****	****	****	****
99	aC-(N)CO	0.0403	0.3436	0.1155	-18.2987	-2.3892	****	****	****	****	****
100	NHCONH	****	****	0.1766	-20.2163	-2.8684	****	****	****	****	****
101	NH_2CONH	****	****	-0.0339	-20.0323	****	****	****	****	****	****
102	NH_2CON	****	****	0.6960	-19.7493	****	****	****	****	****	****
103	NHCON	-1.1943	****	0.6042	-19.0795	****	****	****	****	****	****
104	NCON	****	-0.6828	0.4337	-17.3382	****	****	****	****	****	****
105	$aC\text{-}NHCONH_2$	****	****	0.0108	-24.1940	****	****	****	****	****	****
106	aC-NHCONH	-1.7196	1.9217	0.1090	-25.0710	-1.1863	****	****	****	****	****
107	NHCO except as above	0.5780	-1.3767	0.1661	-14.2912	0.4049	****	****	****	****	****
108	CH_2Cl	0.6285	-0.4994	0.3771	-17.2028	0.3827	2.1156	0.3421	-0.2757	-1.4515	****
109	CHCl	0.0499	0.5406	0.1587	-17.0769	0.2987	-0.6211	0.6224	****	-1.3742	****
110	CCl	****	****	0.1362	-17.1631	****	****	****	****	****	****
111	$CHCl_2$	0.5537	0.0038	0.2677	-28.9652	0.4192	1.5563	1.2588	-0.0233	-1.7873	****
112	CCl_2	1.2087	****	-0.0255	-28.5311	1.0330	****	****	****	****	****
113	CCl_3	1.4243	0.8069	0.3228	-41.4755	1.0662	1.6771	2.7478	1.8173	-0.0399	-0.0132
114	CH_2F	****	****	2.1471	-10.6675	****	****	1.8130	0.8584	****	****
115	CHF	****	****	****	-9.3840	****	****	****	-0.6900	****	****
116	CF	****	****	****	****	****	****	****	****	****	****
117	CHF_2	0.8937	****	0.3330	-17.3008	****	****	1.4605	1.6870	-0.5628	****
118	CF_2	0.1228	****	0.1107	-18.6620	0.6414	-0.5715	****	-0.0026	-0.2492	****
119	CF_3	0.5746	-0.4514	1.1768	-23.8614	0.5289	1.2195	0.7594	2.1289	0.0883	****
120	CCl_2F	****	****	-0.2444	-35.8860	0.9305	0.3580	2.4605	2.1575	0.0131	0.1546
121	HCClF	****	****	-0.5171	-23.3119	****	-0.7500	1.9636	0.9439	-1.4911	****
122	$CClF_2$	****	****	0.0285	-29.8871	0.6942	0.5700	2.4605	2.0786	-0.0409	****
123	aC-Cl	0.6311	0.4746	0.1892	-17.0037	0.5836	1.0971	0.6906	****	-0.2832	-0.0290
124	aC-F	0.3314	0.0741	0.2803	-10.8936	0.6116	****	0.3160	****	****	****
125	aC-I	0.9468	****	0.2509	-47.8151	0.7213	****	****	****	****	****
126	aC-Br	0.8007	0.7128	0.4868	-32.1342	0.4420	****	****	****	****	****
127	-I except as above	0.9023	-0.2683	0.4364	-43.5393	****	2.9991	****	-2.1289	****	****
128	-Br except as above	0.6833	0.2834	0.4352	-27.4476	0.6304	1.4794	0.9350	1.0645	0.4320	****
129	-F except as above	0.2140	-0.3367	0.1492	-6.7874	0.0237	1.9703	0.5273	1.9617	-0.0074	****
130	-Cl except as above	0.5444	0.0418	-0.0421	-12.3720	0.4939	2.3117	1.4774	1.2757	-0.0056	-0.0682
131	CHNOH	0.6580	****	0.2879	-15.3753	****	****	****	****	****	****
132	CNOH	-0.3905	****	0.1104	-14.8064	-1.2495	****	****	****	****	****

133	aC-CHNOH	1.3584	****	0.9215	****	****	****	****	****	****	****
134	OCH_2CH_2OH	-1.2180	-1.8262	-0.1391	-19.4346	-0.5730	1.7806	0.0004	****	****	****
135	$OCHCH_2OH$	****	****	0.0735	-19.4893	****	-0.6850	-0.1526	****	****	****
136	OCH_2CHOH	0.0290	****	-0.0840	-19.6137	****	0.8671	-0.1808	****	****	****
137	-O-OH	****	****	-0.0096	-9.9157	****	****	0.3052	****	****	****
138	CH_2SH	1.8331	2.4088	0.3595	-16.8160	-2.1222	3.4375	****	****	****	****
139	CHSH	0.2695	****	0.5330	-17.4570	****	****	****	****	****	****
140	CSH	****	****	-0.1180	-16.5064	****	****	****	****	****	****
141	aC-SH	****	****	0.7935	-16.4699	****	3.7519	****	****	****	****
142	-SH except as above	0.5489	****	-0.4044	-10.7037	****	2.6753	****	****	****	****
143	CH_3S	0.2836	2.1003	0.4241	-16.2229	0.9628	2.6895	****	****	****	****
144	CH_2S	0.5397	-1.4511	0.3870	-16.2453	-0.5093	****	****	****	****	****
145	CHS	0.7313	9.4995	0.5624	-16.1516	-0.7805	-1.2947	****	****	****	****
146	CS	1.1701	****	0.3257	-16.5664	****	****	****	****	****	****
147	aC-S-	1.0083	-0.5150	0.4494	-15.7488	-1.0166	****	****	****	****	****
148	SO	-1.5826	-2.4432	0.1023	-14.7151	-1.5532	-0.2469	-0.5536	****	****	****
149	SO2	2.5425	****	0.3801	-21.7306	-1.0301	2.3321	****	****	****	****
150	SO_3 (sulfite)	****	****	-0.1636	-31.4604	****	****	****	****	****	****
151	SO_3 (Sulfonate)	0.4513	****	0.3305	-27.2267	****	****	****	****	****	****
152	SO_4 (Sulfate)	****	****	0.8082	-32.5350	****	2.9901	****	****	****	****
153	aC-SO	-0.5034	****	0.7318	-21.6327	-2.4918	****	****	****	****	****
154	aC-SO_2	0.4065	-0.8851	0.0966	-26.3450	-2.0960	****	****	****	****	****
155	PH (phosphine)	****	****	****	****	****	****	****	****	****	****
156	P (Phospine)	****	****	0.6540	****	****	****	****	****	****	****
157	PO_3 (Phospite)	****	****	0.5089	****	-1.6943	****	****	****	****	****
158	PHO_3 (Phosponate)	0.3773	****	0.0938	-25.8155	****	****	****	****	****	****
159	PO_3 (Phosponate)	0.0507	2.3731	1.3895	-26.0493	-2.9814	****	****	****	****	****
160	PHO_4 (Phospate)	1.5332	****	-0.0158	-30.8958	-1.7167	****	****	****	****	****
161	PO_4 (Phospate)	0.0473	2.3370	1.8277	-30.9346	-2.0276	****	****	****	****	****
162	aC-PO_4	2.0153	4.5391	1.1477	-36.6047	-1.2336	0.1450	****	****	****	****
163	aC-P	****	****	0.6623	****	****	****	****	****	****	****
164	CO_3 (Carbonate)	****	****	0.4013	-20.0014	****	****	1.5293	****	****	****
165	C_2H_3O	0.8118	****	0.3448	-14.1859	****	2.0487	0.7572	****	****	****
166	C_2H_2O	-1.4431	****	0.0268	-14.0802	****	****	****	****	****	****
167	C_2HO	****	****	0.2419	-15.2773	****	****	****	****	****	****
168	CH_2 (cyclic)	0.1038	-0.2352	0.0255	-4.8771	0.2427	0.2430	0.0436	****	****	****
169	CH (cyclic)	0.3421	1.1919	0.3515	-5.0340	-0.2328	-0.1306	0.0555	****	****	****
170	C (cyclic)	-0.6996	0.6314	0.3590	-4.5355	-0.0624	-2.1325	0.0103	****	****	****
171	CH=CH (cyclic)	0.2634	0.5840	0.0512	-9.1129	0.6403	0.9258	-0.1761	****	****	****
172	CH=C (cyclic)	0.9929	-2.9859	0.2542	-9.0957	-1.0636	-1.0568	-0.6613	****	****	****
173	C=C (cyclic)	1.4317	0.2367	0.3198	-9.3336	-0.2606	-1.2451	****	****	****	****
174	CH_2=C (cyclic)	1.1079	2.1640	0.6046	-8.6263	0.1438	-2.4611	-0.6120	****	****	****
175	NH (cyclic)	-0.5856	0.3105	0.1081	-4.9664	-0.4746	3.6453	****	****	****	****
176	N (cyclic)	-0.5091	-3.7536	0.1248	-4.4756	0.1110	-0.1597	-0.1781	****	****	****
177	CH=N (cyclic)	0.0220	0.7125	-0.0103	-9.3002	0.3716	****	****	****	****	****
178	C=N (cyclic)	0.1731	2.6967	0.1735	-9.3647	-0.0603	****	****	****	****	****

179	O (cyclic)	-0.3445	-0.3108	0.0003	-5.0128	-0.7104	1.2330	0.1823	****	****	****
180	CO (cyclic)	-0.2591	-0.8055	0.0617	-9.7579	-0.1690	1.9059	0.3317	****	****	****
181	S (cyclic)	1.3798	-1.2604	0.0539	-11.0518	0.1784	-3.3757	****	****	****	****
182	SO_2(cyclic)	****	0.0804	0.2151	-21.4596	-0.5954	****	****	****	****	****
183	>NH	0.0556	-0.4808	0.3294	-4.6200	-0.5666	****	****	****	****	****
184	-O-	-0.8862	1.7786	0.5619	-5.1116	-0.5432	-6.3984	****	****	****	****
185	-S-	1.0294	2.4264	0.1568	-11.7368	-0.3619	****	****	****	****	****
186	>CO	1.3751	0.8472	0.0875	-9.6930	0.4503	0.5076	0.1421	****	****	****
187	PO_2	****	****	0.1658	****	****	****	****	****	****	****
188	CH-N	****	2.3606	-0.1676	-9.1669	****	****	****	****	****	****
189	SiHO	****	****	****	****	****	****	****	****	****	****
190	SiO	****	****	-0.0142	-16.1963	-0.3192	-2.9881	****	****	****	****
191	SiH_2	****	****	-0.2242	****	****	****	****	****	****	****
192	SiH	****	****	****	****	****	****	****	****	****	****
193	Si	****	****	0.3379	-11.4006	-1.4109	****	****	****	****	****
194	$(CH_3)_3N$	****	****	****	****	****	****	****	****	****	****
195	N=N	0.7152	****	0.0892	-9.4458	****	****	****	****	****	****
196	C_{cyclic}=N-	1.1742	****	0.5802	-9.4880	-0.6167	****	****	****	****	****
197	C_{cyclic}=CH-	1.9779	1.0355	0.6675	-8.9669	-0.6699	0.9004	****	****	****	****
198	C_{cyclic}=NH	****	****	-0.0374	-9.9059	****	****	****	****	****	****
199	N=O	-0.0312	0.7543	0.3193	-9.8426	-0.2232	****	****	****	****	-0.0775
200	C_{cyclic}=C	1.2298	****	0.4824	-9.0043	-1.2043	****	0.0712	****	****	****
201	P=O	1.1226	-0.5256	0.8567	-16.1074	-1.4127	****	****	****	****	****
202	N=N	****	1.1355	0.3544	-10.8546	-0.2014	****	****	****	****	****
203	C=NH	-0.7564	****	-0.3121	-10.0237	-0.6413	****	****	****	****	****
204	>C=S	0.8375	-0.3582	0.2675	-15.7517	0.0854	****	****	****	****	****
205	aC-CON	0.3392	-2.9838	0.2853	-17.8813	-2.7177	****	****	****	****	****
206	aC=O	-0.0010	****	-0.1490	-10.0219	-0.4153	****	****	****	****	****
207	aN-	-0.3370	0.6885	0.2069	-4.2632	-1.1185	****	****	****	****	****
208	-Na	****	****	****	****	****	****	****	****	****	****
209	-K	****	****	****	****	****	****	****	****	****	****
210	HCONH	****	****	-0.1282	-14.4140	-1.0365	****	****	****	****	****
211	CHOCH	1.3384	0.5499	0.8732	-14.0900	****	****	****	****	****	****
212	C_2O	****	****	****	****	****	****	****	****	****	****
213	SiH_3	****	****	****	****	****	****	****	****	****	****
214	SiH_2O	****	****	****	****	****	****	****	****	****	****
215	CH=C=CH	****	****	****	-13.3858	****	****	****	****	****	****
216	CH=C=C	****	****	****	****	****	****	****	****	****	****
217	OP(=S)O	1.0213	3.8226	1.4105	-33.4664	0.0143	0.7109	****	****	****	****
218	R	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
219	$CF_{2cyclic}$	****	****	0.3368	-17.7706	****	****	****	1.0010	****	****
220	CF_{cyclic}	****	****	-0.1246	-11.4927	****	****	****	****	****	****

[a] The symbols $LC50(FM)_{1i}$, $LC50(DM)_{1i}$, $LD50_{1i}$, $LogWs_{1i}$, $BCF_{1i,}$ $PEL_{1i,}$ PCO_{1i}, GWP_{1i}, ODP_{1i},and AP_{1i} represent the contributions (C_i) of the first-order groups for the corresponding properties.

Table B5 (continued). MG Method Based Property Models Analysed Using Simultaneous Regression Method: First-Order Groups and their Contributions [a] for the Properties÷ EUA_C, EUA_{NC}, ERA_C, ERA_{NC}, EFW_C, EFW_{NC}, ESW_C, ESW_{NC}, ENS_C, ENS_{NC}, EAS_C, and EAS_{NC}.

	Group	$EUA_{C\,1i}$	$EUA_{NC\,1i}$	$ERA_{C\,1i}$	$ERA_{NC\,1i}$	$EFW_{C\,1i}$	$EFW_{NC\,1i}$	$ESW_{C\,1i}$	$ESW_{NC\,1i}$	$ENS_{C\,1i}$	$ENS_{NC\,1i}$	$EAS_{C\,1i}$	$EAS_{NC\,1i}$
1	CH_3	0.2768	-0.3889	0.4186	-0.8784	0.7393	-0.3816	-0.3653	-1.0175	0.6318	0.1850	0.6353	0.2561
2	CH_2	-0.1334	-0.0407	-0.1321	-0.0319	-0.1459	-0.0700	-0.1772	-0.0575	0.0849	0.1648	0.0411	0.1062
3	CH	0.0802	0.0707	0.3152	0.5123	0.1677	0.2100	1.0114	0.3785	0.3251	-0.2407	-0.0273	-0.5895
4	C	0.9788	0.9223	0.2656	1.7157	-0.3090	-0.0359	1.9303	0.3055	0.2978	-0.9176	-0.6505	-1.1619
5	CH_2=CH	-0.3402	-1.3987	0.8962	-2.0681	0.9410	-1.6118	-0.2764	-1.1415	0.9785	-0.5469	1.1985	0.0661
6	CH=CH	-1.3145	-2.1775	1.0385	-0.5744	0.4864	-1.2323	0.2899	-0.6143	0.5527	-1.8356	0.6069	-0.7348
7	CH_2=C	-1.3393	-1.3567	1.6881	-0.3947	1.1356	-1.0631	0.7675	0.2210	1.9580	-1.1219	2.0215	-0.0034
8	CH=C	-3.0812	-1.8284	0.2217	-0.8211	-0.4232	-1.1149	0.2755	0.2933	0.2432	-3.2576	0.1952	-1.7401
9	C=C	-3.0673	-2.3768	1.4520	-0.3112	-1.5124	-1.1376	0.8643	1.5679	-1.1132	-4.2931	-1.3559	-2.0206
10	CH_2=C=CH	****	****	****	****	****	****	****	****	****	****	****	****
11	CH_2=C=C	****	****	****	****	****	****	****	****	****	****	****	****
12	C=C=C	****	****	****	****	****	****	****	****	****	****	****	****
13	CH≡C	-4.0744	-1.7053	-0.4115	-2.0220	-0.1562	-1.8383	-0.5146	-1.4962	0.1162	-0.6296	-0.2172	-0.6574
14	C≡C	****	****	****	****	****	****	****	****	****	****	****	****
15	aCH	0.1318	-0.0824	0.1728	-0.3183	0.2732	-0.1185	-0.0798	-0.2644	0.2984	0.1383	0.3088	0.1538
16	aC fused with aromatic ring	-0.1343	0.3895	-0.5355	0.5155	-0.5066	0.5060	-0.4842	0.3773	-0.1470	0.5247	-0.3689	2.2520
17	aC fused with non-aromatic ring	-0.7234	0.1826	-1.0425	0.6396	-0.9302	0.7453	-1.0106	0.5475	-0.8140	0.5110	-0.9460	1.9490
18	aC except as above	-0.8413	-0.0646	-0.7811	0.6090	-1.2107	-0.2465	-1.0751	-0.1323	-0.5215	0.0358	-0.5195	-0.4922
19	aN in aromatic ring	0.0545	-0.5575	-0.1527	-0.8743	0.2922	-0.4505	0.0351	-0.5408	0.1579	-0.3233	0.0617	-0.2065
20	aC-CH_3	-0.5357	-0.0847	0.1353	0.0871	-0.1590	0.1514	-0.5494	0.0327	-0.0284	0.6381	0.0242	0.7085
21	aC-CH_2	-0.5009	0.9054	-0.9434	1.1561	-1.0858	0.4515	-0.6868	0.8390	-0.9057	0.2123	-0.7848	0.1967
22	aC-CH	-0.6200	0.0515	-1.2974	0.9275	-2.2675	-0.5408	-0.3939	-0.1009	-1.4835	-0.8336	-1.4515	-1.1191
23	aC-C	-3.3208	0.2718	-3.9393	1.5295	-6.1374	-0.5875	-3.4758	-0.0850	-5.2596	-1.3713	-5.3546	-1.8340
24	aC-CH=CH_2	0.1127	0.0416	0.2302	0.5380	1.4500	1.4195	-0.1057	0.7205	0.6742	1.6855	1.0702	1.8907
25	aC-CH=CH	1.0394	0.2028	1.9524	-1.0455	1.2025	-0.2276	1.4654	-0.1484	1.6964	-0.7658	1.5229	-1.7432
26	aC-C=CH_2	****	****	****	****	****	****	****	****	****	****	****	****
27	aC-C≡CH	****	****	****	****	****	****	****	****	****	****	****	****
28	aC-C≡C	****	****	****	****	****	****	****	****	****	****	****	****
29	OH	0.4194	0.5362	0.2433	-0.8277	0.8887	0.3100	0.2217	0.5225	0.3945	0.3581	0.6619	0.6345
30	aC-OH	0.5318	-0.1766	0.4224	-0.6903	0.5713	-0.4110	0.7130	0.3440	0.9767	0.4142	1.0735	0.3160
31	COOH	0.9436	-0.3737	0.9872	-1.1003	1.5695	-0.0513	1.0489	0.7155	1.0456	0.8499	0.9371	0.2633
32	aC-COOH	-1.9180	-0.1662	0.6256	-0.9179	1.1534	0.1616	1.2207	1.9706	1.3280	0.2411	1.2017	0.1223
33	CH_3CO	0.0905	1.3373	-0.6548	-0.1061	****	1.2797	-0.5174	0.8365	-0.0497	2.0562	****	2.1799
34	CH_2CO	-3.7621	****	-3.7891	****	-3.6144	****	0.8647	****	-4.7374	****	-4.5274	****
35	CHCO	****	-0.1163	****	0.1359	****	-0.3423	****	0.2564	****	-0.4092	****	-1.0036
36	CCO	****	****	****	****	****	****	****	****	****	****	****	****
37	aC-CO	-3.6693	0.1230	-3.2999	-0.2251	-2.1671	-0.3041	-2.1170	0.0400	-3.3091	-0.0111	-3.0206	-0.4405
38	CHO	0.7415	-1.0946	0.3433	-1.4468	0.7471	-0.8811	-0.3255	-0.4580	0.4532	0.2603	0.5964	0.1867
39	aC-CHO	1.5855	-0.1253	1.1053	-0.3669	1.1466	-0.2715	1.0434	-0.1487	0.8989	0.6655	0.9465	0.3748

40	CH_3COO	0.4512	0.9230	0.4035	0.0309	0.6553	1.1335	0.1280	0.2915	0.4063	1.7726	0.5420	1.4246
41	CH_2COO	1.1595	1.2031	0.1357	0.7204	0.4287	1.4075	1.5236	2.3643	-0.9216	1.2035	-0.0509	1.0455
42	CHCOO	****	0.3293	****	1.0144	****	2.4473	****	2.5891	****	0.6452	****	0.1441
43	CCOO	0.1221	****	-0.2906	****	-0.8719	****	0.3265	****	-0.5178	****	-0.8982	****
44	HCOO	****	****	****	****	****	****	****	****	****	****	****	****
45	aC-COO	-0.4819	0.4459	-0.4124	0.4075	-0.4451	0.5003	-0.1062	1.3013	-0.2463	0.4928	-0.4509	-0.0204
46	aC-OOCH	****	****	****	****	****	****	****	****	****	****	****	****
47	aC-OOC	****	0.5722	2.5567	0.2878	-0.7154	0.5013	6.7362	1.0671	1.0188	0.2805	1.1582	-0.0771
48	COO except as above	0.2977	-1.5617	-0.3523	-1.4325	-0.1730	-1.8607	-0.1199	-0.8076	-0.1119	-0.3508	-0.0340	-0.0515
49	CH_3O	0.6256	0.0811	-0.2026	-0.9191	0.6929	-0.0290	-0.7151	-0.2724	0.3522	0.3233	0.5985	0.5144
50	CH_2O	0.0748	0.1975	0.4215	-0.2890	0.1084	0.1516	-1.0241	0.1937	0.2143	-0.6273	0.1613	-0.3999
51	CH-O	****	-0.0821	****	-0.4530	****	-0.6497	****	0.5274	****	-1.2816	****	-1.7349
52	C-O	-5.2568	****	0.9916	****	-0.5381	****	-3.1651	****	-1.4749	****	-2.4229	****
53	aC-O	-0.0802	-0.2859	-0.0713	-0.0853	-0.4911	-0.2147	0.2796	0.6630	-0.3339	-0.9133	-0.2773	-0.8960
54	CH_2NH_2	2.6349	****	2.1445	****	2.7445	****	2.2313	****	2.2150	****	2.5740	****
55	$CHNH_2$	-0.6207	****	-0.4900	****	-0.4117	****	-0.1089	****	-0.9221	****	-0.6954	****
56	CNH_2	****	****	****	****	****	****	****	****	****	****	****	****
57	CH_3NH	0.6012	3.5226	1.1697	0.8357	0.5009	3.0770	0.0504	3.3927	0.5130	4.1468	0.6086	2.6084
58	CH_2NH	-0.0418	1.1184	0.2209	1.8720	-0.2442	****	0.2238	-0.4997	0.0384	****	0.0274	****
59	CHNH	****	-1.6773	****	-1.5280	****	-1.6357	****	-1.2850	****	-3.0506	****	-3.2766
60	CH_3N	-0.6436	0.9099	-0.8553	-0.1924	-0.8106	0.3857	-0.6728	1.2439	-0.6269	-0.2579	-0.5892	0.0841
61	CH_2N	0.4475	-2.0098	0.0993	-2.4756	-0.3946	-2.7768	0.6252	-2.9607	-0.2504	-2.2740	-0.3272	-5.6684
62	aC-NH_2	0.4078	-0.0630	0.3147	-0.3145	0.2777	-0.2788	0.1703	-0.0925	0.3313	-0.0620	0.5484	0.0665
63	aC-NH	-0.4107	0.4817	-0.7748	0.3599	-1.1879	0.0024	-0.3339	0.7840	-0.8853	-0.2743	-0.8825	-0.4021
64	aC-N	-0.4531	1.8888	-0.3214	2.3247	-1.6764	1.1305	-0.0422	1.7549	-1.2858	-0.0682	-1.5082	-0.5018
65	NH_2 except as above	-0.3084	-0.6762	-0.5831	-1.2756	-0.0303	-1.2682	-0.4261	0.0098	-0.1112	0.1261	-0.0458	0.3308
66	CH=N	-0.9032	-0.8004	0.3303	-0.9608	-0.0639	-0.9532	-0.5617	-1.0261	0.2969	-0.0600	0.3659	-0.4586
67	C=N	0.4874	0.7296	0.0018	1.3884	-0.1105	0.1401	0.2676	0.2538	0.8550	0.9988	0.7018	0.1943
68	CH_2CN	****	****	****	****	****	****	****	****	****	****	****	****
69	CHCN	****	****	****	****	****	****	****	****	****	****	****	****
70	CCN	-0.6082	****	-0.8192	****	-1.0567	****	0.2996	****	-1.4462	****	-0.9915	****
71	aC-CN	1.0739	-0.1684	0.9127	-0.4935	1.2596	-0.0803	1.2767	-0.0518	1.2816	0.0372	1.2918	0.2045
72	CN except as above	0.8164	-0.2555	-0.6385	-1.4802	0.7394	-0.2068	0.3631	-0.5179	0.9006	0.1745	0.9294	0.7612
73	CH_2NCO	****	****	****	****	****	****	****	****	****	****	****	****
74	CHNCO	****	****	****	****	****	****	****	****	****	****	****	****
75	CNCO	****	****	****	****	****	****	****	****	****	****	****	****
76	aC-NCO	1.1330	-2.1259	0.4514	-2.3525	0.5267	-2.0876	-0.4091	-2.3986	1.1762	-1.2588	1.1214	-1.4601
77	CH2NO_2	****	****	****	****	****	****	****	****	****	****	****	****
78	$CHNO_2$	-0.7932	-1.2349	-2.0288	-2.3937	-1.2622	-1.0859	-1.9514	-1.5630	-1.6619	-1.1247	-1.2586	-1.1625
79	CNO_2	-0.9724	****	1.4716	****	-0.2284	****	-2.2112	****	1.3183	****	0.7714	****
80	aC-NO_2	0.1245	-0.7880	-0.5036	-1.1595	-0.0264	-0.6468	-0.4483	-0.4066	0.1379	-0.1310	0.3479	-0.1339
81	NO_2 except as above	-0.0683	0.8138	-1.1988	-0.2626	-0.1981	0.8788	-0.5692	0.2233	-0.7637	2.1380	-0.4482	2.0678
82	ONO	0.0475	****	-0.7753	****	0.0152	****	-1.8329	****	-0.3701	****	0.0539	****
83	ONO_2	0.4673	****	-0.0272	****	0.4472	****	-0.1621	****	0.2459	****	0.4794	****
84	$HCON(CH_2)_2$	****	****	****	****	****	****	****	****	****	****	****	****
85	$HCONHCH_2$	****	****	****	****	****	****	****	****	****	****	****	****

86	$CONH_2$	0.6701	-0.6830	0.4481	-0.6628	1.4781	-0.7170	1.9978	-0.3006	1.2097	0.1667	1.3995	0.6316
87	$CONHCH_3$	-0.3275	-1.0584	-1.4288	-2.0736	-0.1886	-1.5198	-0.6454	-1.2248	-1.0779	-0.3723	-0.8224	-0.7812
88	$CONHCH_2$	1.9906	1.8243	0.9711	0.8861	0.7879	2.0946	2.5279	1.7462	1.6381	2.5388	1.6578	2.4492
89	$CON(CH_3)_2$	1.3855	-0.1214	-0.5104	-1.2591	-0.2235	-0.7867	0.1119	-0.5949	-0.9044	0.2034	-0.1317	0.1891
90	$CONCH_3CH_2$	****	****	****	****	****	****	****	****	****	****	****	****
91	$CON(CH_2)_2$	-0.9625	0.2621	-1.5819	0.2178	-2.1596	-0.3609	0.1099	0.4682	-1.5442	-0.7156	-1.3442	-1.2618
92	CONHCO	****	****	****	****	****	****	****	****	****	****	****	****
93	CONCO	****	-0.3656	****	-0.0544	****	-0.2574	****	1.3425	****	-0.6528	****	-0.6126
94	aC-$CONH_2$	1.3190	****	0.4221	****	1.2138	****	0.5586	****	0.6842	****	1.0114	****
95	aC-NH(CO)H	****	****	****	****	****	****	****	****	****	****	****	****
96	aC-N(CO)H	****	****	****	****	****	****	****	****	****	****	****	****
97	aC-CONH	-3.6015	1.3725	-2.4897	1.0603	-1.3548	1.2428	-0.8792	1.3704	-2.5327	0.7981	-1.7480	0.7555
98	aC-NHCO	0.3259	0.3311	0.0376	0.8914	-0.2426	0.7337	0.9831	1.1069	-0.3504	0.6114	-0.0668	0.5521
99	aC-(N)CO	-0.8438	0.8189	-0.9527	1.2103	-2.1512	0.6037	0.1744	2.7438	-0.9767	-0.2882	-0.9029	-0.4164
100	NHCONH	0.9715	****	0.8013	****	1.0687	****	1.6118	****	0.6872	****	0.7296	****
101	NH_2CONH	1.1915	****	-0.0835	****	0.8273	****	0.9456	****	0.5512	****	0.6840	****
102	NH_2CON	0.3622	****	-1.0023	****	-0.3280	****	0.0298	****	-0.2842	****	-0.3256	****
103	NHCON	1.9570	0.5198	0.1885	1.0282	0.7393	1.0205	1.8051	3.8658	0.8960	-1.7350	0.5601	-1.4133
104	NCON	****	****	****	****	****	****	****	****	****	****	****	****
105	aC-$NHCONH_2$	****	****	****	****	****	****	****	****	****	****	****	****
106	aC-NHCONH	****	0.8147	1.6760	1.1664	0.6501	0.4940	****	0.3930	1.4306	-0.1933	1.6286	-0.1363
107	NHCO except as above	-0.1165	-0.6416	-0.4101	-0.1778	-0.1483	-0.5051	0.7002	-0.3826	-0.1430	-0.7959	-0.0158	-0.3768
108	CH_2Cl	-0.3133	-0.9231	-0.3902	-1.5409	-0.0012	-0.8957	-0.7134	-1.7334	-0.0092	-0.3747	0.0485	-0.0864
109	CHCl	0.7229	0.2038	-0.3138	-0.2775	0.4587	-0.4287	-1.0845	-0.5868	0.1700	-0.1056	0.3100	0.1483
110	CCl	****	****	****	****	****	****	****	****	****	****	****	****
111	$CHCl_2$	0.4190	-0.9871	-0.1853	-2.4110	0.6353	-1.0667	-1.0422	-1.6383	0.3711	-0.3263	0.4982	-0.2269
112	CCl_2	****	-0.4318	****	-0.5196	****	-0.6916	****	-0.5364	****	-0.4357	****	-0.2253
113	CCl_3	0.3605	-1.2506	-0.4669	-2.4145	0.2786	-1.5568	-1.8568	-2.2684	-0.1843	-0.6057	0.0910	-0.5476
114	CH_2F	1.7119	0.3666	1.0533	-0.8101	1.8094	0.7354	0.6295	0.0059	1.4211	0.4477	1.6553	0.6039
115	CHF	****	****	****	****	****	****	****	****	****	****	****	****
116	CF	****	****	****	****	****	****	****	****	****	****	****	****
117	CHF_2	****	0.0703	****	-0.4804	****	0.1719	****	-1.1833	****	1.9698	****	1.6025
118	CF_2	****	****	****	****	****	****	****	****	****	****	****	****
119	CF_3	0.9473	0.8895	0.1773	-0.3682	1.0720	0.3204	-1.3467	-1.0247	0.5889	1.5174	0.8026	1.7167
120	CCl_2F	****	-1.4099	****	-2.1593	****	-1.7969	****	-3.1682	****	-1.4242	****	-1.1754
121	HCClF	****	1.7393	****	0.7917	****	1.3080	****	-0.1021	****	1.6004	****	1.9043
122	$CClF_2$	****	0.7293	****	-0.9851	****	0.8903	****	0.0847	****	1.4088	****	1.5158
123	aC-Cl	-0.0445	-0.2890	-0.3069	-0.7079	-0.1654	-0.4874	-0.8314	-0.8213	-0.1399	-0.1050	-0.0633	-0.0999
124	aC-F	-0.9985	-0.2195	-1.2055	-0.4714	-0.7630	-0.2717	-1.4596	0.0464	-0.6826	0.0476	-1.2152	0.2171
125	aC-I	****	****	****	****	****	****	****	****	****	****	****	****
126	aC-Br	0.2093	-0.2519	0.0753	-0.6226	0.3176	-0.2506	-0.0476	-0.4843	0.4777	-0.0766	0.2864	-0.1032
127	-I except as above	****	****	****	****	****	****	****	****	****	****	****	****
128	-Br except as above	-0.1961	-0.3216	-0.1708	-1.2917	0.0232	-0.3045	-0.9789	-1.1443	-0.2296	0.0657	-0.0047	0.3657
129	-F except as above	0.6207	-0.7864	-0.5794	-2.2689	0.6589	-0.5031	-1.9094	-0.9024	-0.1512	0.0691	0.2341	0.1207
130	-Cl except as above	-1.9393	-0.1480	-0.7165	-2.0875	-0.3256	-0.0976	-1.4034	-1.2269	0.0434	0.3658	-0.4433	0.5279
131	CHNOH	****	****	****	****	****	****	****	****	****	****	****	****

132	CNOH	0.1658	****	-1.0768	****	-0.6315	****	0.2332	****	-0.3801	****	-0.0535	****
133	aC-CHNOH	****	****	****	****	****	****	****	****	****	****	****	****
134	OCH_2CH_2OH	2.3891	-0.0244	1.7788	-1.2602	2.1537	-0.4130	2.4413	0.9540	1.6546	0.2117	1.8160	0.3751
135	$OCHCH_2OH$	****	****	****	****	****	****	****	****	****	****	****	****
136	OCH_2CHOH	****	2.4833	****	0.8375	****	2.0173	****	4.0842	****	1.7767	****	1.9349
137	-O-OH	****	****	****	****	****	****	****	****	****	****	****	****
138	CH_2SH	-0.6250	****	-1.9924	****	-1.1190	****	-0.9189	****	-0.0570	****	0.1574	****
139	CHSH	****	****	****	****	****	****	****	****	****	****	****	****
140	CSH	****	****	****	****	****	****	****	****	****	****	****	****
141	aC-SH	-1.2297	****	-3.7114	****	-2.1847	****	-0.4976	****	-1.8963	****	-1.8779	****
142	-SH except as above	****	****	****	****	****	****	****	****	****	****	****	****
143	CH_3S	6.3958	-0.1547	0.9366	-0.3871	1.3187	1.0294	1.6307	-0.3173	1.1446	-0.4689	1.1055	-0.2242
144	CH_2S	-0.3057	-0.0492	-1.0545	-0.3098	-1.3645	-0.2406	-0.6499	-0.3212	-1.0185	-0.2384	-1.0862	-0.3142
145	CHS	****	-1.3648	****	-1.1438	****	-2.1457	****	-3.5270	****	-3.4275	****	-3.2466
146	CS	7.6552	-0.9007	3.4256	3.9642	2.6418	-1.3869	3.8903	-0.5260	1.0593	-2.3698	2.8471	-2.3476
147	aC-S-	****	0.2217	-1.0230	0.1749	****	-0.0818	-1.3298	0.1865	****	-0.2047	-1.0183	-0.2664
148	SO	1.7021	4.3538	0.9399	2.4815	0.3344	4.3129	0.3485	4.3583	-0.3864	3.2094	-0.0328	0.2414
149	SO2	****	-0.3824	****	-0.0907	****	0.0637	****	0.4310	****	1.5538	****	1.1670
150	SO_3 (sulfite)	-0.0827	****	-0.3225	-1.3680	-1.0811	-0.6689	-1.5796	0.8753	-0.7085	****	-0.4129	0.3305
151	SO_3 (Sulfonate)	-0.2533	****	-1.7071	****	-0.9481	****	0.5266	****	-1.0911	****	-0.7997	****
152	SO_4 (Sulfate)	****	****	****	****	****	****	****	****	****	****	****	****
153	aC-SO	****	-0.6036	****	-0.7755	****	-0.7480	****	0.6670	****	-1.0006	****	-0.8386
154	aC-SO_2	-2.3992	-0.2423	-2.8465	-0.3578	-2.9877	-1.4779	-1.6383	-1.5973	-3.2071	-0.1751	-3.0574	-0.8298
155	PH (phosphine)	****	****	****	****	****	****	****	****	****	****	****	****
156	P (Phospine)	****	-1.0674	****	-0.5893	****	-2.1490	****	-1.7547	****	-0.9683	****	-0.7173
157	PO_3 (Phospite)	****	****	****	****	****	****	****	****	****	****	****	****
158	PHO_3 (Phosponate)	0.8052	****	-0.3669	****	-0.0439	****	1.2624	****	-0.1360	****	0.1555	****
159	PO_3 (Phosponate)	1.3267	0.0298	0.7963	-0.0321	-0.4619	0.1061	2.3937	2.0215	0.2316	-0.8312	0.2283	-0.2818
160	PHO_4 (Phospate)	****	****	****	****	****	****	****	****	****	****	****	****
161	PO_4 (Phospate)	1.6283	-0.5208	-0.2120	0.0581	0.3359	-0.6785	2.0237	1.0577	0.5070	-0.5538	0.2061	-1.0503
162	aC-PO_4	****	****	****	****	****	****	****	****	****	****	****	****
163	aC-P	****	****	****	****	****	****	****	****	****	****	****	****
164	CO_3 (Carbonate)	****	****	****	****	****	****	****	****	****	****	****	****
165	C_2H_3O	-0.3133	-5.1385	-1.2673	-4.7155	1.1448	-4.7775	-0.1153	-7.7255	0.0260	-3.6274	0.9980	-4.0098
166	C_2H_2O	****	-1.0203	****	4.2383	****	-1.7704	****	2.1725	****	-1.8586	****	-4.8444
167	C_2HO	****	****	****	****	****	****	****	****	****	****	****	****
168	CH_2 (cyclic)	0.0478	0.0383	0.1037	-0.2351	0.0218	-0.0591	-0.1159	-0.1360	0.2430	0.2778	0.2042	0.2686
169	CH (cyclic)	-0.5705	-1.3636	-0.9411	-0.1917	-0.3534	-1.7395	0.2311	-3.3705	-0.6251	-0.5402	-0.1148	-2.6460
170	C (cyclic)	3.9613	1.0459	1.1634	4.6781	0.4554	0.8751	1.6771	2.9940	-0.9880	0.1288	0.4143	-2.4011
171	CH=CH (cyclic)	-0.2006	-0.3717	-0.1551	-0.8137	-0.3694	-0.4172	-1.1639	-0.7469	0.0672	0.7392	0.0478	0.6207
172	CH=C (cyclic)	0.0562	0.0417	0.2281	-0.4702	-0.4847	-0.2261	-0.1890	-0.5499	0.4875	-0.5866	0.2344	-0.9710
173	C=C (cyclic)	-0.8232	-3.6378	-1.2786	-0.8741	-2.1582	-3.0260	-0.9217	-3.1614	-2.1348	-4.2929	-1.7673	-2.0572
174	CH_2=C (cyclic)	-0.3177	****	0.8420	****	-0.3089	****	-1.0253	****	1.8053	****	1.4118	****
175	NH (cyclic)	0.2068	-0.6063	-0.2440	-1.2240	0.3465	-0.7304	-0.0197	-0.3171	0.2312	-0.6572	0.2839	-0.4248
176	N (cyclic)	0.7465	-0.9135	-0.1568	-0.0631	0.5222	-1.0567	0.5583	-0.0837	-0.0920	-2.0578	0.0099	-1.9278
177	CH=N (cyclic)	0.5309	0.0584	1.4538	-0.5310	0.6452	0.4018	-0.7708	-0.1928	-0.1078	1.0813	0.1624	1.5244

178	C=N (cyclic)	-0.1023	-0.5296	-0.0508	-0.9244	-0.2505	-0.4985	-0.3228	-0.4970	-0.0218	0.2954	-0.0959	0.3652
179	O (cyclic)	-0.6160	-1.3405	-0.4177	-1.5397	0.4395	-1.2790	-0.1866	-1.6388	-0.5459	-1.2752	-0.2245	-0.8744
180	CO (cyclic)	-0.1467	0.7802	0.0273	0.3561	0.0079	0.8369	0.6613	1.1518	0.0450	0.9198	0.0179	0.8191
181	S (cyclic)	-0.0113	-0.4065	-0.3833	-0.2744	0.3451	-0.2271	0.4112	-0.2562	-0.1345	0.1483	0.0751	0.3262
182	SO_2 (cyclic)	0.7597	-0.1628	-0.2927	-1.0910	1.4921	-0.9201	1.4024	-0.9510	4.2417	-1.0633	4.3097	-0.8285
183	>NH	-0.3552	-0.1971	-0.3453	0.2549	-0.1154	-0.8038	-0.2881	-2.1316	-0.1419	-1.3337	-0.1565	-0.9680
184	-O-	2.1320	-1.0893	-0.7573	-1.8584	-0.3473	-1.2017	-3.3651	1.1068	-1.5205	-0.8072	-1.3733	0.1945
185	-S-	0.3962	-0.3454	0.8282	0.0048	-1.0937	-0.6376	0.3392	-0.6979	0.5789	-0.2663	0.2048	-0.1624
186	>CO	****	-0.7055	****	-0.5992	-2.3290	-0.2531	-0.3051	-0.5422	****	-0.2774	-2.3418	-0.4289
187	PO_2	****	****	****	****	****	****	****	****	****	****	****	****
188	CH-N	-0.7970	0.4190	-0.7895	-0.0521	****	-1.0914	****	0.6603	-1.6551	-1.2596	****	-1.9068
189	SiHO	****	****	****	****	****	****	****	****	****	****	****	****
190	SiO	****	****	****	****	****	****	****	****	****	****	****	****
191	SiH_2	****	****	****	****	****	****	****	****	****	****	****	****
192	SiH	****	****	****	****	****	****	****	****	****	****	****	****
193	Si	****	****	****	****	****	****	****	****	****	****	****	****
194	$(CH_3)_3N$	****	****	****	****	****	****	****	****	****	****	****	****
195	N=N	****	-0.6798	****	-1.3422	****	-0.7632	****	-2.2405	****	-0.4732	****	0.7231
196	C_{cyclic}=N-	0.3309	-0.3663	-0.5439	-1.4243	0.1428	0.5389	-0.0601	0.0130	0.6748	0.7104	0.3442	0.9730
197	C_{cyclic}=CH-	-1.0346	****	-0.2734	****	0.3924	****	0.7763	****	-0.4543	****	-0.9067	****
198	C_{cyclic}=NH	****	****	****	****	****	****	****	****	****	****	****	****
199	N=O	-1.8430	****	-1.2750	****	-1.2403	****	-1.3171	****	-1.4893	****	-1.1820	****
200	C_{cyclic}=C	-1.4671	****	-0.8366	****	-0.0991	****	-0.2696	****	0.2699	****	-0.5165	****
201	P=O	-5.6926	-1.8136	-0.5494	-0.9626	-1.0689	-1.7597	-0.2499	-0.7070	-0.9191	-1.2194	-0.9084	-1.9008
202	N=N	0.5131	****	-0.9381	****	-0.3220	****	-0.1820	****	-0.8614	****	-0.9633	****
203	C=NH	****	****	****	****	****	****	****	****	****	****	****	****
204	>C=S	0.4140	-0.4898	0.1229	-0.0780	-0.0878	-0.2683	0.2252	-0.1931	-0.2965	0.7406	-0.1104	0.2244
205	aC-CON	0.7937	-3.7639	-0.8587	0.1618	-1.0781	-4.1159	1.3443	****	-0.6874	-5.3844	-0.8192	-5.1798
206	aC=O	0.6202	****	1.2167	****	1.3160	****	2.2897	****	1.3151	****	1.6017	****
207	aN-	-0.9102	-1.3674	-0.4206	-1.1407	-0.7226	-1.4421	0.1246	-0.6597	-0.2579	-0.9406	0.0202	-2.6866
208	-Na	****	****	****	****	****	****	****	****	****	****	****	****
209	-K	****	****	****	****	****	****	****	****	****	****	****	****
210	HCONH	0.7064	****	0.1018	****	0.8454	****	0.5271	****	0.4107	****	0.5743	****
211	CHOCH	****	****	****	****	****	****	****	****	****	****	****	****
212	C_2O	****	****	****	****	****	****	****	****	****	****	****	****
213	SiH_3	****	****	****	****	****	****	****	****	****	****	****	****
214	SiH_2O	****	****	****	****	****	****	****	****	****	****	****	****
215	CH=C=CH	****	****	****	****	****	****	****	****	****	****	****	****
216	CH=C=C	****	****	****	****	****	****	****	****	****	****	****	****
217	OP(=S)O	0.8795	-0.3584	1.6200	0.0413	0.1379	-1.0782	4.2446	-0.3807	-0.0978	-0.9345	-0.2608	-1.2441
218	R	****	****	****	****	****	****	****	****	****	****	****	****
219	$CF_{2cyclic}$	****	****	****	****	****	****	****	****	****	****	****	****
220	CF_{cyclic}	-1.1229	****	-1.9644	****	-3.4286	****	-3.5188	****	-3.8755	****	-3.4357	****

[a] The symbols $EUA_{C\,1i}$, $EUA_{NC\,1i}$, $ERA_{C\,1i}$, $ERA_{NC\,1i}$, $EFW_{C\,1i}$, $EFW_{NC\,1i}$, $ESW_{C\,1i}$, $ESW_{NC\,1i}$, $ENS_{C\,1i}$, $ENS_{NC\,1i}$, $EAS_{C\,1i}$, and $EAS_{NC\,1i}$ represent the contributions (C_i) of the first-order groups for the corresponding properties.

Table B6. MG Method Based Property Models Analysed Using Simultaneous Regression Method: Second-Order Groups and their Contributions [a] for the Properties÷ LC_{50}(FM), LC_{50}(DM), LD_{50}, $LogW_s$, BCF, PEL(OSHA-TWA), and PCO

	Group	$LC50(FM)_{2j}$	$LC50(DM)_{2j}$	$LD50_{2j}$	$LogWs_{2j}$	BCF_{2j}	PEL_{2j}	$PCO1_{2j}$
1	$(CH_3)_2CH$	-0.0418	0.2712	-0.0144	-0.2681	0.3714	-0.0027	0.0313
2	$(CH_3)_3C$	0.3137	4.1440	0.1404	-0.1698	0.1995	0.0414	0.1580
3	$CH(CH_3)CH(CH_3)$	-0.0995	****	-0.0161	-0.3705	1.2525	-0.0350	0.0765
4	$CH(CH_3)C(CH_3)_2$	****	****	-0.0632	0.3614	1.1101	-0.0102	0.0354
5	$C(CH_3)_2C(CH_3)_2$	****	****	-0.0830	-0.3041	****	-0.0684	0.1280
6	CH_n=CH_m-CH_p=CH_k (*k,m,n,p* in 0..2)	-0.1772	0.8857	0.0942	-0.5589	-1.2538	-0.5349	0.5521
7	CH_3-CH_m=CH_n (*m,n* in 0..2)	0.0749	-3.5587	-0.1077	-0.0577	0.5100	-0.1718	-0.0364
8	CH_2-CH_m=CH_n (*m,n* in 0..2)	-0.2902	2.4471	-0.0401	0.0703	0.0765	0.2436	0.0642
9	CH_p-CH_m=CH_n (*m,n* in 0..2; *p* in 0..1)	0.0477	-0.5971	0.1941	0.0841	-0.0935	****	0.0636
10	CHCHO or CCHO	0.3970	****	0.0140	-0.5415	****	****	-0.0825
11	CH_3COCH_2	0.0279	1.0878	0.0083	0.4242	****	0.4724	-0.4764
12	CH_3COCH or CH_3COC	0.5626	****	0.3218	0.6789	-0.3859	0.3350	-0.2513
13	CHCOOH or CCOOH	-0.2421	****	0.1954	0.2890	0.1200	0.4992	-0.1341
14	CH_3COOCH or CH_3COOC	0.0097	****	0.1502	0.0716	****	-0.0526	-0.0621
15	CO-O-CO	****	****	0.0176	-0.6806	0.0000	1.6800	-10000.0000
16	CHOH	-0.0975	3.9109	0.1155	0.2588	-0.2924	0.1197	0.0487
17	COH	0.0935	****	0.2278	0.8332	-0.4952	-0.2124	0.4127
18	CH_3COCH_nOH (*n* in 0..2)	****	****	0.0054	0.9579	****	****	-0.5912
19	NCCHOH or NCCOH	****	****	0.2700	0.5755	****	0.0000	-10000.0000
20	OH-CH_n-COO (*n* in 0..2)	1.3209	****	0.0627	0.1891	1.1097	****	-0.4055
21	$CH_m(OH)CH_n(OH)$ (*m,n* in 0..2)	-1.3676	0.0361	0.1196	-0.4326	-0.7323	0.2582	-0.0929
22	$CH_m(OH)CH_n(NH_p)$ (*m,n,p* in 0..2)	-0.4231	1.2993	-0.0387	0.1678	0.0000	0.0860	-0.0417
23	$CH_m(NH_2)CH_n(NH_2)$ (*m,n* in 0..2)	-0.1456	-3.8727	-0.5152	-0.8213	****	-1.9649	****
24	$CH_m(NH)CH_n(NH_2)$ (*m,n* in 1..2)	****	****	-0.3115	0.3349	0.0000	-1.3100	-10000.0000
25	$H_2NCOCH_nCH_mCONH_2$ (*m,n* in 1..2)	****	****	****	-0.7592	****	****	****
26	$CH_m(NH_n)$-COOH (*m,n* in 0..2)	1.2919	5.1623	-0.1667	-0.4295	****	****	****
27	HOOC-CH_n-COOH (*n* in 1..2)	****	****	-0.0211	0.4420	****	****	****
28	HOOC-CH_n-CH_m-COOH (*n, m* in 1..2)	****	****	0.1245	0.1019	3.0655	****	-0.4783
29	HO-CH_n-COOH (*n* in 1..2)	****	****	-0.0243	-0.4726	0.8399	****	-0.0784
30	NH_2-CH_n-CH_m-COOH (*n, m* in 1..2)	****	****	****	-0.7592	****	****	****
31	CH_3-O-CH_n-COOH (n in 1..2)	****	****	****	0.8111	****	****	****
32	HS-CH-COOH	0.0000	****	0.7610	0.8831	-10000.0000	-1.6100	-10000.0000
33	HS-CH_n-CH_m-COOH (*n, m* in 1..2)	****	****	0.0428	1.5568	****	****	****
34	NC-CH_n-CH_m-CN (*n, m* in 1..2)	****	****	0.7607	0.5459	****	-1.5017	****

35	OH-CH_n-CH_m-CN (*n, m* in 1..2)	****	****	0.3340	1.0784	****	****	****
36	HS-CH_n-CH_m-SH (*n, m* in 1..2)	****	****	-2.7796	****	****	****	****
37	COO-CH_n-CH_m-OOC (*n, m* in 1..2)	****	****	0.1410	-0.4872	****	****	-0.2074
38	OOC-CH_m-CH_m-COO (*n, m* in 1..2)	-0.2344	-16.0297	-0.0563	0.3977	0.0000	0.0000	0.5500
39	NC-CH_n-COO (*n* in 1..2)	****	****	-0.3421	-0.2022	****	****	****
40	$COCH_nCOO$ (*n* in 1..2)	****	****	-0.0528	0.2418	****	****	****
41	CH_m-O-CH_n=CH_p (*m,n,p* in 0..3)	****	****	-0.4559	0.2041	1.6032	****	****
42	CH_m=CH_n-F (*m,n* in 0..2)	****	****	-0.1574	0.0268	****	-1.2982	****
43	CH_m=CH_n-Br (*m,n* in 0..2)	****	0.0974	-0.3767	0.4918	****	0.7909	****
44	CH_m=CH_n-I (*m,n* in 0..2)	****	****	****	0.0987	****	****	****
45	CH_m=CH_n-Cl (*m,n* in 0..2)	-0.2955	0.2971	0.1633	0.2219	-0.0569	-1.7406	-0.9410
46	CH_m=CH_n-CN (*m,n* in 0..2)	0.2974	4.3101	0.2436	0.9999	****	0.1129	0.0000
47	CH_n=CH_m-COO-CH_p (*m,n,p* in 0..3)	-0.1220	-0.0227	-0.1420	0.0383	****	0.0206	0.0788
48	CH_m=CH_n-CHO (*m,n* in 0..2)	1.8419	9.3811	0.1660	0.2322	****	0.0772	0.3566
49	CH_m=CH_n-COOH (*m,n* in 0..2)	1.5593	****	-0.0256	-0.0141	****	-1.0172	-0.3647
50	aC-CH_n-X (*n* in 1..2) X: Halogen	-0.1264	****	-0.0660	0.1818	-0.5132	0.6493	****
51	aC-CH_n-NH_m (*n* in 1..2; *m* in 0..2))	-0.5979	****	-0.2187	0.7638	-0.1295	****	****
52	aC-CH_n-O- (*n* in 1..2)	0.2287	-2.8691	0.0546	0.4210	0.1076	****	****
53	aC-CH_n-OH (*n* in 1..2)	0.0090	****	-0.0002	0.2106	-0.3066	****	-0.1689
54	aC-CH_n-CN (*n* in 1..2)	2.2929	-2.2058	0.0990	0.5638	0.1066	****	****
55	aC-CH_n-CHO (*n* in 1..2)	****	****	-0.3084	0.4866	****	****	****
56	aC-CH_n-SH (*n* in 1..2)	****	****	0.6393	****	****	****	****
57	aC-CH_n-COOH (*n* in 1..2)	****	****	0.3223	0.2576	-0.6363	****	****
58	aC-CH_n-CO- (*n* in 1..2)	0.0000	****	0.1190	0.3883	-0.3370	-10000.0000	-10000.0000
59	aC-CH_n-S- (*n* in 1..2)	****	****	-0.5222	0.1038	1.1115	****	****
60	aC-CH_n-OOC-H (*n* in 1..2)	****	****	0.2836	****	****	****	****
61	aC-CH_m-NO_2 (*n* in 1..2)	****	****	****	****	****	****	****
62	aC-CH_n-$CONH_2$ (*n* in 1..2)	****	****	0.1814	0.5403	****	****	****
63	aC-CH_n-OOC (*n* in 1..2)	-0.4720	-0.4810	-0.1659	-0.7164	-0.3023	0.7772	****
64	aC-CH_n-COO (*n* in 1..2)	-0.6542	-6.1400	0.0539	-0.1301	0.0458	****	****
65	aC-SO_2-OH	-0.5382	****	-0.1870	1.0300	1.5586	****	****
66	aC-$CH(CH_3)_2$	0.0851	-1.2681	-0.3378	-0.2936	0.1272	-1.1271	0.0187
67	aC-$C(CH_3)_3$	0.9465	****	-0.1434	0.7323	-0.0775	5.6444	0.0907
68	aC-CF_3	0.1166	0.3533	-0.7788	1.9593	1.6118	****	-0.2101
69	$(CH_n{=}C)_{cyc}$-CHO (*n* in 0..2)	****	****	-0.1031	-0.5395	****	0.2158	****
70	$(CH_n{=}C)_{cyc}$-COO-CH_m (*n,m* in 0..3)	****	****	-0.2405	-0.0249	0.7906	****	****
71	$(CH_n{=}C)_{cyc}$-CO- (*n* in 0..2)	-0.6127	****	0.1205	-0.2214	****	****	****
72	$(CH_n{=}C)_{cyc}$-CH_3 (*n* in 0..2)	-0.3382	5.1943	-0.1284	0.1409	-0.2985	-2.5355	-0.2129
73	$(CH_n{=}C)_{cyc}$-CH_2 (*n* in 0..2)	-0.8382	-2.2216	-0.0219	0.4412	****	0.1846	****
74	$(CH_n{=}C)_{cyc}$-CN (*n* in 0..2)	-0.0901	****	-0.6381	-0.3809	****	****	****

75	$(CH_n{=}C)_{cye}$-Cl (*n* in 0..2)	-0.2396	0.4435	0.0291	0.1416	-0.3723	-0.7029	****
76	CH_{cye}-CH_3	0.0321	0.7237	-0.2120	0.1410	0.1321	-0.3707	-0.0854
77	CH_{cye}-CH_2	-0.5631	1.4904	-0.2032	0.3006	0.7223	-1.1768	-0.0314
78	CH_{cye}-CH	0.6735	-0.5530	-0.2439	0.3390	-0.0178	****	0.0269
79	CH_{cye}-C	-0.6977	****	-0.4669	0.0614	1.3638	****	-0.0579
80	CH_{cye}-CH=CH_n (*n* in 1..2)	0.0547	-3.3258	-0.0845	-0.0140	0.4312	1.4915	****
81	CH_{cye}-C=CH_n (*n* in 1..2)	-0.6174	2.9140	-0.4441	0.5030	-2.8314	****	0.8562
82	CH_{cye}-Cl	-0.3183	-0.9800	-0.0649	0.2801	0.2869	-1.2358	****
83	CH_{cye}-F	****	****	****	0.5092	****	****	****
84	CH_{cye}-OH	-0.2028	1.0683	-0.1594	-0.0298	-0.2648	0.0492	-0.1266
85	CH_{cye}-NH_2	****	****	-0.4730	0.1626	****	0.2963	****
86	CH_{cye}-NH-CH_n (*n* in 0..3)	0.1644	-0.7336	-0.3636	0.6737	0.0436	****	****
87	CH_{cye}-N-CH_n (*n* in 0..3)	1.2144	****	-0.4944	0.6127	0.8210	****	****
88	CH_{cye}-SH	****	****	1.6279	****	****	0.9254	****
89	CH_{cye}-CN	****	****	-0.5241	0.9679	****	****	****
90	CH_{cye}-COOH	0.3786	****	-0.3532	0.0041	0.9628	****	****
91	CH_{cye}-CO	****	****	-0.0387	0.4932	****	****	****
92	CH_{cye}-NO_2	****	****	0.4482	0.9901	****	****	****
93	CH_{cye}-S-	-1.6152	****	-0.2856	0.8968	****	****	****
94	CH_{cye}-CHO	0.5061	****	-0.3566	-0.0628	****	****	-0.0370
95	CH_{cye}-O-	****	****	-0.4260	0.2118	****	2.1204	****
96	CH_{cye}-OOCH	****	****	0.0509	****	****	****	****
97	CH_{cye}-COO	-0.7603	-1.1127	-0.4650	-0.0666	0.2877	****	****
98	CH_{cye}-OOC	-0.4578	-2.0921	-0.3600	0.2191	****	****	****
99	C_{cye}-CH_3	0.5272	-0.6499	-0.1205	-0.0642	-0.0631	2.4390	-0.0002
100	C_{cye}-CH_2	1.2910	-0.7346	-0.0003	0.4062	-0.5229	1.0713	-0.0319
101	C_{cye}-OH	-0.2373	-1.9432	0.2718	-0.3541	-1.6360	-1.1715	****
102	$>N_{cye}$-CH_3	0.5632	6.9066	0.0506	0.4460	-1.2764	****	-0.2283
103	$>N_{cye}$-CH_2	-0.3114	1.9548	0.0864	0.5720	-0.5023	****	****
104	AROMRINGs^{1}s^2	0.2621	0.2164	0.0716	-0.1196	0.1105	-0.4639	-0.2439
105	AROMRINGs^{1}s^3	-0.1180	0.1443	0.0941	-0.2204	0.0452	-0.5725	-0.2979
106	AROMRINGs^{1}s^4	0.2146	0.3769	0.0225	-0.1809	-0.0630	0.0847	-0.1949
107	AROMRINGs^{1}s^{2}s^3	-0.0720	0.3149	0.1488	0.0218	0.3649	-0.6955	-0.2645
108	AROMRINGs^{1}s^{2}s^4	0.2048	0.0338	0.0290	-0.0516	0.1049	-0.9649	-0.1603
109	AROMRINGs^{1}s^{3}s^5	-0.0500	-0.1186	0.0747	-0.3272	0.4939	-0.7364	-0.2885
110	AROMRINGs^{1}s^{2}s^{3}s^4	0.4147	-0.4579	-0.1170	0.1183	0.0951	****	-0.1422
111	AROMRINGs^{1}s^{2}s^{3}s^5	-0.0011	-0.0506	-0.0488	0.0218	0.2253	-3.3441	-0.0812
112	AROMRINGs^{1}s^{2}s^{4}s^5	0.3494	-0.5777	0.0164	-0.2460	0.5129	****	-0.1437
113	PYRIDINEs^2	-0.9133	****	0.4753	0.9191	-0.1885	****	****
114	PYRIDINEs^3	-0.4916	****	-0.0360	0.9089	-1.0214	****	****

115	PYRIDINEs^{4}	-0.7575	****	0.2708	0.4137	-0.6056	****	****
116	PYRIDINEs^{2}s^{3}	-1.1257	****	0.5297	0.4975	****	****	****
117	PYRIDINEs^{2}s^{4}	****	****	0.2268	0.2720	****	****	****
118	PYRIDINEs^{2}s^{5}	-0.6158	0.2969	0.0882	0.2547	1.5410	****	****
119	PYRIDINEs^{2}s^{6}	-0.2151	****	0.3627	-0.0288	****	****	****
120	PYRIDINEs^{3}s^{4}	****	****	0.2749	0.1079	****	****	****
121	PYRIDINEs^{3}s^{5}	****	****	-0.2183	-0.1755	****	****	****
122	PYRIDINEs^{2}s^{3}s^{6}	****	****	-0.5987	-0.2032	****	****	****
123	(CH^{n}=CH^{m})^{cyc}-COOH	****	****	-0.9943	-0.5168	****	****	****
124	AROMRINGs^{1}s^{2}s^{3}s^{4}s^{5}	-0.0481	-0.6732	0.1845	0.3517	0.1277	****	0.0474
125	aC-NHCOCH$_2$N	****	****	0.6326	1.1063	****	****	****
126	(N=C)$_{cyc}$-CH$_3$	-0.5150	****	0.0283	0.6955	-2.0489	****	****
127	aC-CONH(CH$_2$)$_2$N	****	****	0.2557	0.7687	****	****	****
128	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	-0.1618	0.2409	0.7943	****	****
129	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	0.2385	0.8555	****	****	****
130	aC-SO$_2$NH$_n$ (n>=0;n<3)	****	****	-0.3188	0.2319	1.6555	****	****

[a] The symbols LC50(FM)$_{2j}$, LC50(DM)$_{2j}$LD50$_{2j}$, LogWs$_{2j}$, BCF$_{2j,}$ PEL$_{2j,}$ PCO$_{2j}$ represent the contributions (D_j) of the second-order groups for the corresponding properties. Note that there are no second-order contributions for GWP, ODP and AP since these properties belong to chemicals whose structure is defined by first-order groups alone.

Table B6 (continued). MG Method Based Property Models Analysed Using Simultaneous Regression Method: Second-Order Groups and their Contributions [a] for the Properties÷ EUA_C, EUA_{NC}, ERA_C, ERA_{NC}, EFW_C, EFW_{NC}, ESW_C, ESW_{NC}, ENS_C, ENS_{NC}, EAS_C, and EAS_{NC}.

	Group	$EUA_{C\,2j}$	$EUA_{NC\,2j}$	$ERA_{C\,2j}$	$ERA_{NC\,2j}$	$EFW_{C\,2j}$	$EFW_{NC\,2j}$	$ESW_{C\,2j}$	$ESW_{NC\,2j}$	$ENS_{C\,2j}$	$ENS_{NC\,2j}$	$EAS_{C\,2j}$	$EAS_{NC\,2j}$	$EUA_{C\,2j}$
1	$(CH_3)_2CH$	0.4454	0.7777	0.0850	0.9533	-0.0271	0.6037	0.3571	1.0021	-0.1402	0.9090	0.2321	1.0531	0.4454
2	$(CH_3)_3C$	-1.0261	0.7628	-0.3870	0.2392	-0.2564	0.6369	-1.2769	1.2133	-0.8796	1.1569	0.1557	0.4723	-1.0261
3	$CH(CH_3)CH(CH_3)$	****	****	****	****	****	****	****	****	****	****	****	****	****
4	$CH(CH_3)C(CH_3)_2$	****	****	****	****	****	****	****	****	****	****	****	****	****
5	$C(CH_3)_2C(CH_3)_2$	****	****	****	****	****	****	****	****	****	****	****	****	****
6	CH_n=CH_m-CH_p=CH_k (*k,m,n,p* in 0..2)	1.7943	2.7047	-0.9739	0.2460	-0.9041	1.5568	-0.7469	1.3600	-0.2300	2.2724	-0.4340	0.4036	1.7943
7	CH_3-CH_m=CH_n (*m,n* in 0..2)	0.8992	1.1266	-0.0583	0.2893	-0.1388	0.5668	0.0073	0.1956	-0.1366	1.3767	-0.0679	0.5363	0.8992
8	CH_2-CH_m=CH_n (*m,n* in 0..2)	0.6538	0.2386	-0.3203	1.2598	-0.4937	0.5536	-0.9016	-0.0152	-0.3931	0.7228	-0.6658	0.1489	0.6538
9	CH_p-CH_m=CH_n (*m,n* in 0..2; *p* in 0..1)	-0.1003	-1.0487	-1.2161	-1.0125	-0.9001	-1.3775	-1.0900	-1.1375	-1.0261	-1.2026	-1.3403	-1.3735	-0.1003
10	CHCHO or CCHO	****	****	****	****	****	****	****	****	****	****	****	****	****
11	CH_3COCH_2	****	-0.7929	****	0.0988	0.2147	-0.4836	****	-0.2261	****	-0.4645	0.0150	-0.6226	****
12	CH_3COCH or CH_3COC	****	****	****	****	****	****	****	****	****	****	****	****	****
13	CHCOOH or CCOOH	-0.3208	0.0291	-1.4573	-0.5071	-1.4573	-0.3280	-1.5177	-0.5655	-1.2028	-0.7276	-0.8287	-0.2615	-0.3208
14	CH_3COOCH or CH_3COOC	-0.9094	****	-0.1570	****	-0.3162	****	-1.2374	****	-1.1378	****	-0.5791	****	-0.9094
15	CO-O-CO	****	****	****	****	****	****	****	****	****	****	****	****	****
16	CHOH	-0.5119	-0.8836	-0.6160	0.1893	-1.0520	-0.8619	-0.6700	-0.9111	-0.4799	-0.2670	-0.5091	-0.3797	-0.5119
17	COH	-0.3245	-2.3976	-1.5910	-2.1498	-1.7123	-1.5682	-1.1338	-1.8517	-1.0148	-1.1666	-1.2208	-1.0103	-0.3245
18	CH_3COCH_nOH (*n* in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
19	NCCHOH or NCCOH	****	****	****	****	****	****	****	****	****	****	****	****	****
20	OH-CH_n-COO (*n* in 0..2)	-0.2640	-2.5789	-1.3986	-2.2692	-4.3790	-4.1296	-1.7941	-4.8709	-3.8361	1.2855	-2.3116	1.1969	-0.2640
21	$CH_m(OH)CH_n(OH)$ (*m,n* in 0..2)	-2.6993	****	****	-0.4078	****	****	****	****	****	****	****	****	-2.6993
22	$CH_m(OH)CH_n(NH_p)$ (*m,n,p* in 0..2)	-1.6155	****	-1.8221	****	-2.0086	****	-2.5453	****	-2.2315	****	-2.2681	****	-1.6155
23	$CH_m(NH_2)CH_n(NH_2)$ (*m,n* in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
24	$CH_m(NH)CH_n(NH_2)$ (*m,n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
25	$H_2NCOCH_nCH_mCONH_2$ (*m,n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
26	$CH_m(NH_n)$-COOH (*m,n* in 0..2)	-0.1495	****	-0.3633	****	-0.4518	0.7218	-0.6857	****	-0.1520	0.3571	0.1525	1.3200	-0.1495
27	HOOC-CH_n-COOH (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
28	HOOC-CH_n-CH_m-COOH (*n, m* in	****	****	****	****	****	****	****	****	****	****	****	****	****
29	HO-CH_n-COOH (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
30	NH_2-CH_n-CH_m-COOH (*n, m* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
31	CH_3-O-CH_n-COOH (n in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
32	HS-CH-COOH	****	****	****	****	****	****	****	****	****	****	****	****	****
33	HS-CH_n-CH_m-COOH (*n, m* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
34	NC-CH_n-CH_m-CN (*n, m* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****

35	OH-CH_n-CH_m-CN (*n, m* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
36	HS-CH_n-CH_m-SH (*n, m* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
37	COO-CH_n-CH_m-OOC (*n, m* in 1..2)	0.4211	****	1.1314	****	1.2644	****	-0.3272	****	1.5168	****	0.9057	****	0.4211
38	OOC-CH_m-CH_m-COO (*n, m* in 1..2)	-1.5241	3.6672	-0.9983	3.6117	-0.6690	4.6673	-1.0778	5.4329	1.4117	4.1810	0.0326	3.9405	-1.5241
39	NC-CH_n-COO (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
40	$COCH_n$COO (*n* in 1..2)	-2.5340	****	-2.0342	****	-2.6213	****	-4.7273	****	0.0760	****	-1.0854	****	-2.5340
41	CH_m-O-CH_n=CH_p (*m,n,p* in 0..3)	****	****	****	****	****	****	****	****	****	****	****	****	****
42	CH_m=CH_n-F (*m,n* in 0..2)	0.3786	****	0.1935	****	-0.0139	****	1.4655	****	0.6120	****	0.3624	****	0.3786
43	CH_m=CH_n-Br (*m,n* in 0..2)	****	0.6806	****	0.5414	****	0.2753	****	0.0771	****	0.9150	****	0.3664	****
44	CH_m=CH_n-I (*m,n* in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
45	CH_m=CH_n-Cl (*m,n* in 0..2)	2.8096	0.4492	0.1697	1.2437	0.9242	0.0103	0.5423	0.0245	0.3108	0.6621	0.9509	-0.0168	2.8096
46	CH_m=CH_n-CN (*m,n* in 0..2)	-0.2510	-0.7132	-0.6328	-0.4505	-1.2833	-0.4240	-2.0346	-1.5374	-1.6324	-0.4531	-1.7620	-1.4200	-0.2510
47	CH_n=CH_m-COO-CH_p (*m,n,p* in 0..3)	0.4041	1.7879	-0.9669	0.9356	-0.5581	1.7600	0.4096	1.5587	-0.5862	1.3524	-0.4495	0.7187	0.4041
48	CH_m=CH_n-CHO (*m,n* in 0..2)	-0.2428	-1.7213	-1.0203	-1.7625	-0.8167	-0.6646	0.0942	-3.1005	-0.9783	-1.7167	-1.0652	-1.9796	-0.2428
49	CH_m=CH_n-COOH (*m,n* in 0..2)	-1.0794	****	-2.7056	****	-2.6653	****	-2.8203	****	-2.2235	****	-1.9522	****	-1.0794
50	aC-CH_n-X (*n* in 1..2) X: Halogen	2.3361	****	0.8709	****	1.3713	****	0.9893	****	0.6633	****	1.2260	****	2.3361
51	aC-CH_n-NH_m (*n* in 1..2; *m* in 0..2))	1.0687	****	0.6693	****	0.6495	****	0.5118	****	0.9552	****	0.9981	****	1.0687
52	aC-CH_n-O- (*n* in 1..2)	6.0723	0.8650	0.7509	0.8463	2.0927	0.8340	8.1599	0.5995	3.2998	1.1806	2.5796	1.3480	6.0723
53	aC-CH_n-OH (*n* in 1..2)	0.7158	-1.0641	1.2554	-0.4911	1.2648	-1.0102	1.4622	-0.7395	1.1325	-1.3372	1.0486	-1.3613	0.7158
54	aC-CH_n-CN (*n* in 1..2)	****	0.6234	****	0.7483	****	0.5938	****	0.8119	****	0.1023	****	-0.5529	****
55	aC-CH_n-CHO (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
56	aC-CH_n-SH (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
57	aC-CH_n-COOH (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
58	aC-CH_n-CO- (*n* in 1..2)	****	-0.4253	****	-0.3207	****	0.0434	****	0.2368	****	-0.7374	****	-0.4952	****
59	aC-CH_n-S- (*n* in 1..2)	****	-0.4363	****	-0.6230	****	0.1549	****	0.2885	****	1.0150	****	0.9068	****
60	aC-CH_n-OOC-H (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
61	aC-CH_m-NO_2 (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
62	aC-CH_n-$CONH_2$ (*n* in 1..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
63	aC-CH_n-OOC (*n* in 1..2)	2.1857	-0.8637	1.5944	-0.5155	1.7073	-0.2397	2.2699	-0.3791	1.8470	-0.4877	1.5479	-0.5280	2.1857
64	aC-CH_n-COO (*n* in 1..2)	0.6946	2.4109	1.1951	1.7731	3.0189	3.2424	1.4124	3.0979	2.4991	0.0813	1.5682	-0.2439	0.6946
65	aC-SO_2-OH	****	****	****	****	****	****	****	****	****	****	****	****	****
66	aC-$CH(CH_3)_2$	-0.6750	0.3543	-1.5063	-0.2398	-0.9091	0.7285	0.6240	1.5970	-0.6537	-0.2616	-0.4466	0.4541	-0.6750
67	aC-$C(CH_3)_3$	3.0869	1.3842	2.6282	****	3.8946	****	3.3901	****	3.7734	3.3323	3.7719	****	3.0869
68	aC-CF_3	2.7512	1.2533	5.7320	4.7670	4.2528	1.3657	8.4099	1.9079	5.3555	1.0867	4.3307	1.3526	2.7512
69	$(CH_n{=}C)_{cyc}$-CHO (*n* in 0..2)	2.1717	1.2393	0.8882	1.6531	0.8121	0.8593	1.6104	2.0521	1.2094	0.5196	1.0755	0.8454	2.1717
70	$(CH_n{=}C)_{cyc}$-COO-CH_m (*n,m* in 0..3)	****	6.2390	****	2.1891	****	4.7572	****	5.4373	****	2.2894	****	-0.4511	****
71	$(CH_n{=}C)_{cyc}$-CO- (*n* in 0..2)	****	3.6588	****	1.5130	****	1.7279	****	2.6167	****	1.7446	****	0.2172	****
72	$(CH_n{=}C)_{cyc}$-CH_3 (*n* in 0..2)	0.0144	1.6249	0.2798	0.8405	0.6716	1.7910	-0.2938	2.6849	-0.1270	2.3612	0.1740	1.0475	0.0144
73	$(CH_n{=}C)_{cyc}$-CH_2 (*n* in 0..2)	0.8733	-0.4514	0.1148	0.5525	0.0473	-0.3333	0.5312	-0.0406	-0.5548	1.1360	-0.4881	0.8778	0.8733
74	$(CH_n{=}C)_{cyc}$-CN (*n* in 0..2)	****	2.8009	****	1.0755	****	1.6933	****	1.5060	****	2.8493	****	0.0293	****

75	$(CH_n{=}C)_{cyc}$-Cl (*n* in 0..2)	2.4853	0.8650	1.3227	1.0251	1.5691	0.5254	1.5760	1.1605	1.3515	1.0003	1.6352	0.3695	2.4853
76	CH_{cyc}-CH_3	1.3286	3.5106	0.6868	1.6797	-0.4296	3.1837	-0.2258	5.4130	0.1520	2.6395	-0.3265	3.2632	1.3286
77	CH_{cyc}-CH_2	0.8847	3.8741	1.1317	2.1603	-0.3996	3.4800	-0.4124	6.1122	0.0876	3.1659	-0.5379	3.5261	0.8847
78	CH_{cyc}-CH	****	-0.9238	****	2.3200	****	0.7457	****	6.1796	****	0.7722	****	2.8086	****
79	CH_{cyc}-C	****	****	****	****	****	****	****	****	****	****	****	****	****
80	CH_{cyc}-CH=CH_n (*n* in 1..2)	2.3806	1.0425	1.8181	2.9808	0.9041	1.1953	0.2788	3.5141	1.3798	1.1572	0.2457	3.8725	2.3806
81	CH_{cyc}-C=CH_n (*n* in 1..2)	1.1720	2.4730	****	2.6495	-1.7594	2.6481	-0.9170	3.5831	****	0.6485	-2.9697	0.2060	1.1720
82	CH_{cyc}-Cl	2.3759	0.9708	1.2937	1.3585	0.5164	1.1560	0.5041	3.6834	0.4250	-0.1607	0.4479	1.8211	2.3759
83	CH_{cyc}-F	****	****	****	****	****	****	****	****	****	****	****	****	****
84	CH_{cyc}-OH	-1.3384	****	1.3004	****	2.3858	****	-0.9943	****	2.7345	****	0.7397	****	-1.3384
85	CH_{cyc}-NH_2	****	1.6149	****	1.4558	****	2.6148	****	3.5476	****	0.1803	****	2.2142	****
86	CH_{cyc}-NH-CH_n (*n* in 0..3)	****	0.8163	****	-0.7386	****	1.3080	****	3.3778	****	0.6604	****	2.4904	****
87	CH_{cyc}-N-CH_n (*n* in 0..3)	****	****	****	****	****	****	****	****	1.1339	****	****	****	****
88	CH_{cyc}-SH	****	****	****	****	****	****	****	****	****	****	****	****	****
89	CH_{cyc}-CN	****	****	****	****	****	****	****	****	****	****	****	****	****
90	CH_{cyc}-COOH	-0.1192	****	-0.0258	****	-0.8363	****	-0.2251	****	0.1528	****	-0.3452	****	-0.1192
91	CH_{cyc}-CO	-2.0389	****	-0.4304	****	0.8105	****	0.8953	****	0.5810	****	-0.5264	****	-2.0389
92	CH_{cyc}-NO_2	****	****	****	****	****	****	****	****	****	****	****	****	****
93	CH_{cyc}-S-	****	3.3988	****	2.0753	****	4.0643	****	6.2953	****	1.4916	****	3.3534	****
94	CH_{cyc}-CHO	****	****	****	****	****	****	****	****	****	****	****	****	****
95	CH_{cyc}-O-	-2.2400	1.6824	-0.1341	1.9541	-0.6127	2.8569	-1.1417	3.6426	-0.9768	-0.4946	-1.8716	2.9381	-2.2400
96	CH_{cyc}-OOCH	****	****	****	****	****	****	****	****	****	****	****	****	****
97	CH_{cyc}-COO	1.1500	5.8631	2.5531	-0.4490	1.0973	7.4198	3.3882	7.2877	3.2162	3.5543	3.4197	2.2421	1.1500
98	CH_{cyc}-OOC	-2.6029	****	-1.7657	****	0.5728	****	-2.1338	****	-1.5802	****	-2.7917	****	-2.6029
99	C_{cyc}-CH_3	-1.2509	-0.7606	-0.4669	-1.1415	-0.3483	-0.7903	-0.0784	-1.1166	0.3823	-0.8171	-0.2401	1.2917	-1.2509
100	C_{cyc}-CH_2	-2.1932	-1.6180	-1.4721	-2.9398	-1.1779	-1.6672	-1.1572	-3.6152	-0.5213	-0.5739	-1.0688	0.5956	-2.1932
101	C_{cyc}-OH	-4.1529	****	-2.6517	****	-2.0853	****	-2.1961	****	-1.3210	****	-2.4397	****	-4.1529
102	$>N_{cyc}$-CH_3	-2.6236	1.2241	-0.6600	1.0941	-1.6712	1.8657	-0.3848	2.4545	-0.5769	1.1601	-1.1689	3.0465	-2.6236
103	$>N_{cyc}$-CH_2	-2.3077	0.4390	-1.6670	0.3304	-2.0988	0.3703	-5.2922	0.7772	-1.0751	0.4878	-1.1755	-0.0732	-2.3077
104	AROMRINGs^{1}s^2	-0.2932	0.3709	-0.7671	0.2911	-0.5028	0.2541	0.0123	0.0459	-0.5034	0.0076	-0.6684	0.3018	-0.2932
105	AROMRINGs^{1}s^3	-0.0386	-0.2628	0.2833	-0.7480	0.1137	-0.7375	-0.3048	-0.8663	-0.4912	-0.4973	-0.2536	-0.6056	-0.0386
106	AROMRINGs^{1}s^4	0.2368	-0.2471	-0.5227	-0.0877	-0.2714	0.0439	0.2671	0.0037	-0.5270	-0.1168	-0.6095	0.1734	0.2368
107	AROMRINGs^{1}s^{2}s^3	-0.8549	-0.8951	-1.4756	-1.0252	-1.5103	-1.2564	-1.0988	-1.9660	-1.7515	-1.0944	-1.9131	-1.2851	-0.8549
108	AROMRINGs^{1}s^{2}s^4	0.2638	0.1433	-0.2493	-0.5598	-0.0359	-0.4147	0.5094	-0.8571	-0.3209	-0.6798	-0.6706	-0.5102	0.2638
109	AROMRINGs^{1}s^{3}s^5	0.9570	1.0957	-0.1781	-2.1081	-0.0445	1.1302	0.7663	-1.6879	0.2129	1.4766	-0.2338	1.4124	0.9570
110	AROMRINGs^{1}s^{2}s^{3}s^4	-0.4034	0.1981	-0.6652	0.4054	-0.7230	-0.1871	-0.0139	-1.7111	-0.7600	0.2076	-1.0221	0.1758	-0.4034
111	AROMRINGs^{1}s^{2}s^{3}s^5	0.2665	-0.0685	0.1411	-0.1653	0.2443	-0.2365	0.7054	-0.8478	0.3094	-0.1337	-0.0754	-0.1363	0.2665
112	AROMRINGs^{1}s^{2}s^{4}s^5	1.4506	0.1108	0.0440	0.1174	-0.3537	0.1461	0.3393	0.0988	-0.6499	-0.2285	-0.8629	0.4303	1.4506
113	PYRIDINEs^2	-3.9107	****	-1.3864	****	-1.3452	****	-5.7261	****	-0.5775	****	-1.1313	****	-3.9107
114	PYRIDINEs^3	3.9544	****	2.0083	****	1.0925	****	2.1646	****	1.8698	****	1.7514	****	3.9544

115	PYRIDINEs^{4}	3.7069	****	1.5478	****	0.3720	****	1.3554	****	1.2727	****	0.7867	****	3.7069
116	PYRIDINEs^{2}s^{3}	****	****	****	****	****	****	****	****	****	****	****	****	****
117	PYRIDINEs^{2}s^{4}	1.2066	****	0.9353	****	1.1112	****	2.3016	****	1.0532	****	0.9182	****	1.2066
118	PYRIDINEs^{2}s^{5}	****	-0.7277	****	-1.0006	****	0.3373	****	1.9774	****	0.4128	****	0.5314	****
119	PYRIDINEs^{2}s^{6}	****	****	****	****	****	****	****	****	****	****	****	****	****
120	PYRIDINEs^{3}s^{4}	****	****	****	****	****	****	****	****	****	****	****	****	****
121	PYRIDINEs^{3}s^{5}	****	****	****	****	****	****	****	****	****	****	****	****	****
122	PYRIDINEs^{2}s^{3}s^{6}	****	****	****	****	****	****	****	****	****	****	****	****	****
123	(CH^{n}=CH^{m})^{cyc}-COOH	1.6652	****	0.5202	****	0.8885	****	4.8555	****	-1.0278	****	-0.2002	****	1.6652
124	AROMRINGs^{1}s^{2}s^{3}s^{4}s^{5}	-1.8147	-1.2355	-2.6357	-0.5585	-1.7893	0.0961	-2.1346	-0.4804	-5.3012	-0.1426	-3.6115	-0.8818	-1.8147
125	aC-NHCOCH$_2$N	****	****	****	****	****	****	****	****	****	****	****	****	****
126	(N=C)$_{cyc}$-CH$_3$	1.7027	-2.5588	0.5750	3.4910	1.1188	-1.0985	1.9860	-1.5018	0.4178	-0.1138	0.9384	2.1689	1.7027
127	aC-CONH(CH$_2$)$_2$N	****	****	****	****	****	****	****	****	****	****	****	****	****
128	aC-SO$_2$NH$_n$ (n>=0;n<3)	4.6466	0.6155	4.1474	0.4567	4.4159	1.6027	3.7676	3.0591	4.6311	0.6682	4.6749	0.5460	4.6466
129	aC-SO$_2$NH$_n$ (n>=0;n<3)	4.4830	****	1.8210	****	1.7980	****	1.4651	****	1.3212	****	1.4017	****	4.4830
130	aC-SO$_2$NH$_n$ (n>=0;n<3)	4.9480	****	4.1501	****	4.8739	****	3.8738	****	4.8151	****	4.9756	****	4.9480

[a] The symbols $EUA_{C\,2j}$, $EUA_{NC\,2j}$, $ERA_{C\,2j}$, $ERA_{NC\,2j}$, $EFW_{C\,2j}$, $EFW_{NC\,2j}$, $ESW_{C\,2j}$, $ESW_{NC\,2j}$, $ENS_{C\,2j}$, $ENS_{NC\,2j}$, $EAS_{C\,2j}$, and $EAS_{NC\,2j}$ represent the contributions (D_j) of the second-order groups for the corresponding properties.

Table B7. MG Method Based Property Models Analysed Using Simultaneous Regression Method: Third-Order Groups and their Contributions [a] for the Properties÷ LC_{50}(FM), LC_{50}(DM), LD_{50}, $LogW_s$, BCF, PEL(OSHA-TWA), and PCO

	Group	$LC50(FM)_{3k}$	$LC50(DM)_{3k}$	$LD50_{3k}$	$LogWs_{3k}$	BCF_{3k}	PEL_{3k}	PCO_{3k}
1	HOOC-$(CH_n)_m$-COOH (m>2, n in 0..2)	0.9173	****	-0.4207	0.2040	****	-0.9038	-0.3757
2	NH_n-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	****	1.0560	****	****	****
3	NH_2-$(CH_n)_m$-OH (m>2, n in 0..2)	****	****	-0.2468	****	****	****	****
4	OH-$(CH_n)_m$-OH (m>2, n in 0..2)	****	****	0.2831	1.3300	****	****	-0.1802
5	OH-$(CH_p)_k$-O-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****
6	OH-$(CH_p)_k$-S-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****
7	OH-$(CH_p)_k$-NH_x-$(CH_n)_m$-OH (m,k>0; p,n,x in 0..2)	****	****	****	****	****	****	****
8	CH_p-O-$(CH_n)_m$-OH (m>2; n,p in 0..2)	****	****	****	****	****	****	****
9	NH_2-$(CH_n)_m$-NH_2 (m>2; n in 0..2))	-0.8275	****	0.3173	****	****	-0.9748	****
10	NH_k-$(CH_n)_m$-NH_2 (m>2; k in 0..1; n in 0..2)	****	****	****	****	****	****	****
11	SH-$(CH_n)_m$-SH (m>2; n in 0..2)	****	****	****	****	****	****	****
12	NC-$(CH_n)_m$-CN (m>2)	-1.3554	****	1.3010	1.3500	****	-1.1729	****
13	COO-$(CH_n)_m$-OOC (m>2; n in 0..2)	****	****	****	****	****	****	****
14	aC-$(CH_n$=$CH_m)_{cyc}$ (fused rings) (n,m in 0..1)	0.1769	0.1711	0.2622	0.5630	0.2127	-0.6066	0.5918
15	aC-aC (different rings)	0.5827	2.5396	0.1456	-0.6550	0.3719	-0.0226	****
16	aC-CH_{ncyc} (different rings) (n in 0..1)	0.1573	-0.7346	0.0603	0.5110	0.2323	****	****
17	aC-CH_{ncyc} (fused rings) (n in 0..1)	0.1206	1.1726	0.1955	0.3110	0.0839	0.8204	0.0211
18	aC-$(CH_n)_m$-aC (different rings) (m>1; n in 0..2)	-0.4732	****	-0.4598	-1.1200	-3.6829	****	****
19	aC-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	1.5711	1.0355	-0.0306	-0.3286	****	****	****
20	CH_{cyc}-CH_{cyc} (different rings)	****	****	-0.7228	0.7183	****	****	0.2372
21	CH_{cyc}-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****
22	$CH_{multiring}$	0.1112	-1.1600	-0.2905	0.2640	0.7080	0.4625	0.0488
23	$C_{multiring}$	0.4663	-0.3169	-0.0516	0.2870	-0.5604	-0.8159	-0.0500
24	aC-CH_m-aC (different rings) (m in 0..2)	-0.2180	-0.5207	-0.1477	0.3680	0.3005	-2.2183	0.0870
25	aC-$(CH_m$=$CH_n)$-aC (different rings) (m,n in 0..2)	****	-1.4543	-0.2549	-1.1124	0.0332	****	****
26	$(CH_m$=C$)_{cyc}$-CH=CH-$(C$=$CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****
27	$(CH_m$=C$)_{cyc}$-CH_p-$(C$=$CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****
28	aC-CO-aC (different rings)	-0.2216	****	0.3273	0.9580	-0.4462	****	****
29	aC-CH_m-CO-aC (different rings) (m in 0..2)	****	****	-0.3630	-0.7354	0.5276	****	****

30	aC-CO-(C=CH$_n$)$_{cyc}$ (different rings) (n in 0..1)	****	****	-0.0893	-0.6605	****	****	****
31	aC-CO-CO-aC (different rings)	****	****	****	-0.0946	****	****	****
32	aC-CO$_{cyc}$ (fused rings)	0.6771	1.0063	0.0602	0.0311	0.2109	-1.4914	-0.2966
33	aC-CO-(CH$_n$)$_m$-CO-aC (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****
34	aC-CO-CH$_{ncyc}$ (different rings) (n in 0..1)	****	****	****	-0.5784	****	****	****
35	aC-CO-NH$_n$-aC (different rings) (n in 0..1)	0.0254	****	-0.3798	-0.8170	-0.2883	****	****
36	aC-NH$_n$CONH$_m$-aC (different rings) (n,m in 0..1)	2.2685	1.1010	****	0.6880	-0.7634	****	****
37	aC-CO-N$_{cyc}$ (different rings)	****	****	0.6053	-0.8920	****	****	****
38	aC-S$_{cyc}$ (fused rings)	-1.3941	0.6642	0.1930	0.0047	0.2622	-0.4700	****
39	aC-S-aC (different rings)	-2.3824	****	-0.6622	0.0690	-2.1514	****	****
40	aC-PO$_n$-aC (different rings) (n in 0..4)	****	****	****	****	****	****	****
41	aC-SO$_n$-aC (different rings) (n in 1..4)	****	1.2103	-0.0401	0.0305	****	****	****
42	aC-NH$_{ncyc}$ (fused rings) (n in 0..1)	0.5676	0.7420	0.1005	-0.1850	0.3119	0.6063	****
43	aC-NH-aC (different rings)	0.6705	****	-0.0227	-0.4270	0.6307	-0.7435	****
44	aC-(C=N)$_{cyc}$ (different rings)	****	****	-0.1243	0.5607	-1.1970	****	****
45	aC-(N=CH$_n$)$_{cyc}$ (fused rings) (n in 0..1)	-0.3559	-0.4921	0.2448	-0.0985	-0.3981	****	****
46	aC-(CH$_n$=N)$_{cyc}$ (fused rings) (n in 0..1)	****	****	-0.0772	0.3102	****	****	****
47	aC-O-CH$_n$-aC (different rings) (n in 0..2)	****	****	-0.2144	1.0144	****	****	****
48	aC-O-aC (different rings)	0.3988	1.2922	-0.0525	-0.7136	-0.0772	-0.0011	****
49	aC-CH$_n$-O-CH$_m$-aC (different rings) (n,m in 0..2)	****	****	-0.8111	0.2160	****	****	****
50	aC-O$_{cyc}$ (fused rings)	0.1768	-0.3242	0.1886	-0.6070	0.3585	****	****
51	AROM.FUSED[2]	0.3214	1.4255	-0.1320	-0.0759	-0.1167	0.5784	-0.0172
52	AROM.FUSED[2]s^1	0.8251	1.4112	0.2442	0.0980	0.3046	-0.5039	-0.0424
53	AROM.FUSED[2]s^2	1.0648	1.5552	-0.1073	-0.0711	-0.0191	2.0554	0.0075
54	AROM.FUSED[2]s^{2}s^3	0.5430	6.5271	0.3370	0.1059	0.2984	****	0.1336
55	AROM.FUSED[2]s^{1}s^4	0.0865	****	-0.2060	-0.2716	0.1158	****	****
56	AROM.FUSED[2]s^{1}s^2	****	****	0.1782	-0.0567	-1.1589	****	****
57	AROM.FUSED[2]s^{1}s^3	****	****	0.4300	-0.0732	0.3121	****	****
58	AROM.FUSED[3]	1.0422	0.7528	-0.1893	-0.1255	-0.0916	0.7487	****
59	AROM.FUSED[4a]	****	****	-0.3160	-0.6696	-0.0827	0.0083	****
60	AROM.FUSED[4a]s^1	****	****	-0.5015	0.2968	0.0928	****	****
61	AROM.FUSED[4a]s^{1}s^4	****	****	1.1864	0.3025	****	****	****
62	AROM.FUSED[4p]	2.2573	2.2734	-0.2774	0.0500	-0.2365	0.0083	****
63	AROM.FUSED[4p]s^{3}s^4	****	****	0.2710	2.1472	****	****	****
64	PYRIDINE.FUSED[2]	0.2293	0.8970	0.0213	0.5540	-1.5994	****	****

65	PYRIDINE.FUSED[2-iso]	****	1.0799	0.6366	1.3326	****	****	****
66	PYRIDINE.FUSED[4]	1.4995	2.5289	0.1623	2.0700	-0.5698	-1.0718	****
67	aC-N-CH_{cyc} (different rings)	****	****	0.3013	-0.1996	****	****	****
68	N multiring	-0.0623	6.7397	0.0393	0.4788	****	****	0.1011
69	N_{cyc}-$(CH_2)_3$-N_{cyc} (different rings)	****	****	-0.1340	-0.3802	****	****	****
70	aC-$COCH_2CH_2$-aC (different rings)	****	****	-0.2681	-0.5406	****	****	****
71	aC-O-$(CH_2)_2$-N_{cyc} (different rings)	****	****	-0.6543	****	****	****	****
72	aC-CH(OH)$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****
73	Ncyc-$(CH_2)_2$-CH_{cyc} (different rings)	****	****	-0.3193	****	****	****	****
74	aC-$CONHCH_2$-CH_{cyc} (different rings)	****	****	0.4260	****	****	****	****

[a] The symbols $LC50(FM)_{3k}$, $LC50(DM)_{3k}$, $LD50_{3k}$, $LogWS_{3k}$, $BCF_{3k,}$ $PEL_{3k,}$ PCO_{3k} represent the contributions (E_k) of the third-order groups for the corresponding properties. Note that there are no third-order contributions for GWP, ODP and AP since these properties belong to chemicals whose structure is defined by first-order groups alone.

Table B7 (continued). MG Method Based Property Models Analysed Using Simultaneous Regression Method: Third-Order Groups and their Contributions [a] for the Properties÷ EUA_C, EUA_{NC}, ERA_C, ERA_{NC}, EFW_C, EFW_{NC}, ESW_C, ESW_{NC}, ENS_C, ENS_{NC}, EAS_C, and EAS_{NC}.

	Group	$EUA_{C\ 3k}$	EUA_{NC3k}	$ERA_{C\ 3k}$	ERA_{NC3k}	$EFW_{C\ 3k}$	$EFW_{NC\ 3k}$	$ESW_{C\ 2j}$	$ESW_{NC\ 3k}$	$ENS_{C\ 3k}$	$ENS_{NC\ 3k}$	$EAS_{C\ 3k}$	$EAS_{NC\ 3k}$	$EUA_{C\ 3k}$
1	HOOC-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
2	NH_n-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
3	NH_2-$(CH_n)_m$-OH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
4	OH-$(CH_n)_m$-OH (m>2, n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
5	OH-$(CH_p)_k$-O-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
6	OH-$(CH_p)_k$-S-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
7	OH-$(CH_p)_k$.NH_x-$(CH_n)_m$-OH (m,k>0; p,n,x in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
8	CH_p-O-$(CH_n)_m$-OH (m>2; n,p in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
9	NH_2-$(CH_n)_m$-NH_2 (m>2; n in 0..2))	****	****	****	****	****	****	****	****	****	****	****	****	****
10	NH_k-$(CH_n)_m$-NH_2 (m>2; k in 0..1; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
11	SH-$(CH_n)_m$.SH (m>2; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
12	NC-$(CH_n)_m$-CN (m>2)	****	****	****	****	****	****	****	****	****	****	****	****	****
13	COO-$(CH_n)_m$-OOC (m>2; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
14	aC-$(CH_n{=}CH_m)_{cyc}$ (fused rings) (n,m in 0..1)	1.1453	****	0.3743	****	0.7155	****	1.1082	****	1.1576	****	1.1447	****	1.1453
15	aC-aC (different rings)	-0.3745	0.3502	0.0337	-0.6127	-0.2123	-0.5070	-0.2114	-0.5153	-0.1751	-0.1023	0.2783	0.4820	-0.3745
16	aC-CH_{ncyc} (different rings) (n in 0..1)	-1.6763	1.2288	-0.5860	0.2367	-0.3896	1.8325	-0.5974	2.1670	-0.3506	-0.1667	-1.0244	2.0064	-1.6763
17	aC-CH_{ncyc} (fused rings) (n in 0..1)	0.5989	0.2239	0.4792	0.1614	0.7203	-0.0600	0.8116	-0.3012	0.4990	0.1098	0.4442	0.4050	0.5989
18	aC-$(CH_n)_m$-aC (different rings) (m>1; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
19	aC-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
20	CH_{cyc}-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
21	CH_{cyc}-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
22	$CH_{multiring}$	0.2918	1.4253	0.5641	0.1846	-0.2737	1.5142	-0.8597	3.2326	0.3179	0.5498	-0.2432	2.5758	0.2918
23	$C_{multiring}$	-2.3245	-1.6313	-0.8597	-3.5534	-0.4698	-1.5278	-0.8958	-2.7675	0.8845	-1.0496	-0.0873	1.9041	-2.3245
24	aC-CH_m-aC (different rings) (m in 0..2)	-0.2959	0.9218	0.3732	0.4699	0.3818	1.1639	0.0487	1.1364	1.0739	1.2047	0.9533	1.3399	-0.2959
25	aC-$(CH_m{=}CH_n)$-aC (different rings) (m,n in 0..2)	-1.1581	****	-3.5109	****	-1.3118	****	-2.4897	****	-0.6754	****	-0.7573	****	-1.1581
26	$(CH_m{=}C)_{cyc}$-CH=CH-$(C{=}CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
27	$(CH_m{=}C)_{cyc}$-CH_p-$(C{=}CH_n)_{cyc}$ (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
28	aC-CO-aC (different rings)	2.9341	****	****	****	****	****	2.3524	****	2.2595	****	****	****	2.9341
29	aC-CH_m-CO-aC (different rings) (m in 0..2)	****	****	1.8412	****	1.6105	****	****	****	****	****	2.1609	****	****
30	aC-CO-$(C{=}CH_n)_{cyc}$ (different rings) (n in 0..1)	****	****	****	****	****	****	****	****	****	****	****	****	****
31	aC-CO-CO-aC (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****

32	aC-CO_{cyc} (fused rings)	0.5575	0.2980	0.4436	-0.2414	0.5009	-0.0452	0.1402	0.2220	0.4805	0.2377	0.4643	0.4622	0.5575
33	aC-CO-$(CH_n)_m$-CO-aC (different rings) (m>0; n in	****	****	****	****	****	****	****	****	****	****	****	****	****
34	aC-CO-CH_{ncyc} (different rings) (n in 0..1)	****	****	****	****	****	****	****	****	****	****	****	****	****
35	aC-CO-NH_n-aC (different rings) (n in 0..1)	****	****	****	****	****	****	****	0.9066	****	****	****	****	****
36	aC-NH_nCONH_m-aC (different rings) (n,m in 0..1)	0.8217	****	****	****	****	****	1.5104	****	****	****	****	****	0.8217
37	aC-CO-N_{cyc}(different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
38	aC-S_{cyc} (fused rings)	0.6815	0.6537	1.0950	0.3350	0.8548	1.2864	0.3803	-0.0805	2.1013	-0.4347	1.9715	-0.0939	0.6815
39	aC-S-aC (different rings)	-1.5222	****	****	****	-1.4878	****	****	****	-0.8046	****	****	****	-1.5222
40	aC-PO_n-aC (different rings) (n in 0..4)	****	****	****	****	****	****	****	****	****	****	****	****	****
41	aC-SO_n-aC (different rings) (n in 1..4)	1.2240	****	1.8319	****	2.2857	****	1.8029	****	2.3910	****	2.0459	****	1.2240
42	aC-NH_{ncyc}(fused rings) (n in 0..1)	1.4988	0.4368	1.3033	0.4905	0.9940	0.3361	1.5455	-0.6428	1.6993	0.6135	1.4767	0.6416	1.4988
43	aC-NH-aC (different rings)	4.2975	0.6868	3.7434	1.0236	3.7193	0.5016	3.0226	0.2959	3.5508	0.4208	3.2482	0.6133	4.2975
44	aC-$(C=N)_{cyc}$ (different rings)	1.0499	0.4529	****	-1.7816	-2.6381	0.6892	****	-1.3698	-4.0912	-0.3454	****	0.3300	1.0499
45	aC-$(N=CH_n)_{cyc}$ (fused rings) (n in 0..1)	-0.8892	0.9142	-0.4406	0.9818	0.2535	-0.2085	1.6274	1.4884	-0.0745	-0.0211	0.1517	-0.1760	-0.8892
46	aC-$(CH_n=N)_{cyc}$ (fused rings) (n in 0..1)	****	****	-0.6373	****	****	****	1.5519	****	****	****	-2.1856	****	****
47	aC-O-CH_n-aC (different rings) (n in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
48	aC-O-aC (different rings)	0.2963	0.0230	0.0970	0.5422	0.0342	-0.0853	-0.1136	-0.8360	0.2707	1.6559	0.6668	1.6485	0.2963
49	aC-CH_n-O-CH_m-aC (different rings) (n,m in 0..2)	****	****	****	****	****	****	****	****	****	****	****	****	****
50	aC-O_{cyc} (fused rings)	0.9721	0.8889	1.2643	-0.6400	0.3761	0.1663	0.7579	0.5447	0.9779	0.6311	0.7324	-1.8948	0.9721
51	AROM.FUSED[2]	-0.0316	-0.5677	0.5201	-0.6876	0.1126	-1.1200	0.1693	-0.6070	-0.0125	-0.7338	0.0912	-4.2858	-0.0316
52	AROM.FUSED[2]s^1	0.0476	0.8444	-0.0543	1.2897	-0.2536	0.9398	0.2944	0.6790	0.0195	0.6832	0.0904	1.2549	0.0476
53	AROM.FUSED[2]s^2	0.3853	-0.1511	-0.4057	0.3161	-0.1155	-0.3720	0.1988	-0.4414	0.3824	-0.4358	0.2054	-0.7488	0.3853
54	AROM.FUSED[2]s^2s^3	-3.2511	-1.5652	0.1564	1.6397	-0.6531	-1.2387	-2.7482	-0.9316	-2.5448	0.0525	-2.2312	-0.0208	-3.2511
55	AROM.FUSED[2]s^1s^4	0.0712	****	0.4869	****	0.1806	****	0.7572	****	0.3580	****	0.2650	****	0.0712
56	AROM.FUSED[2]s^1s^2	0.6821	-3.9117	-0.5138	-2.6314	-0.2806	-2.8689	1.3343	-5.4893	1.0399	-0.2754	0.7327	-4.8866	0.6821
57	AROM.FUSED[2]s^1s^3	-1.3732	****	-1.1393	****	-0.6221	****	-0.8382	****	-1.5324	****	-1.8774	****	-1.3732
58	AROM.FUSED[3]	-0.2936	-0.5401	0.5174	-0.5452	0.1298	-1.2250	-0.1360	-1.0354	-0.0145	-0.7079	0.5637	-4.3457	-0.2936
59	AROM.FUSED[4a]	0.1157	****	0.4646	1.0427	0.1454	0.5616	0.8427	0.1918	-0.4812	0.0101	0.2781	0.0172	0.1157
60	AROM.FUSED[4a]s^1	1.0468	****	1.8126	****	1.6412	****	1.9925	****	1.0138	****	1.4593	****	1.0468
61	AROM.FUSED[4a]s^1s^4	-0.7641	****	-2.4896	****	-2.0882	****	-1.8129	****	-2.0344	****	-1.8514	****	-0.7641
62	AROM.FUSED[4p]	-0.5768	-0.4101	-0.4202	-0.6491	-0.0859	-0.5568	0.2636	-0.4714	-0.5330	-0.4819	-0.6143	-2.2533	-0.5768
63	AROM.FUSED[4p]s^3s^4	3.7567	****	0.1636	****	-6.4758	****	7.5539	****	-4.8061	****	-2.2659	****	3.7567
64	PYRIDINE.FUSED[2]	-0.7884	****	-0.5325	****	-0.6751	****	0.4602	****	-0.6388	****	-0.2095	****	-0.7884
65	PYRIDINE.FUSED[2-iso]	****	****	****	****	****	****	****	****	****	****	****	****	****
66	PYRIDINE.FUSED[4]	****	****	****	****	****	****	****	****	****	****	****	****	****
67	aC-N-CH_{cyc} (different rings)	3.4700	****	1.2220	****	-0.1785	****	-2.0455	****	****	****	0.8371	****	3.4700
68	N multiring	0.0626	0.0942	0.2739	-0.9523	-0.5110	0.8827	-0.0914	0.4870	-0.0101	1.4318	0.2564	1.5434	0.0626

69	N_{cyc}-$(CH_2)_3$-N_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
70	aC-$COCH_2CH_2$-aC (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
71	aC-O-$(CH_2)_2$-N_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
72	aC-CH(OH)$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
73	Ncyc-$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****
74	aC-$CONHCH_2$-CH_{cyc} (different rings)	****	****	****	****	****	****	****	****	****	****	****	****	****

[a] The symbols $EUA_{C\ 3k}$, $EUA_{NC\ 3k}$, $ERA_{C\ 3k}$, $ERA_{NC\ 3k}$, $EFW_{C\ 3k}$, $EFW_{NC\ 3k}$, $ESW_{C\ 3k}$, $ESW_{NC\ 3k}$, $ENS_{C\ 3k}$, $ENS_{NC\ 3k}$, $EAS_{C\ 3k}$, and $EAS_{NC\ 3k}$ represent the contributions (E_k) of the third-order groups for the corresponding properties.

Table B8. CI Method Based Property Models: Atom Contributions and Model Constants for the Properties÷ LC_{50}(FM), LC_{50}(DM), LD_{50}, $LogW_s$, BCF, PEL(OSHA-TWA), PCO, GWP, ODP, and AP.

Parameter	LC_{50} (FM)	LC_{50} (DM)	LD_{50}	$LogW_s$	BCF	PEL	PCO	GWP	ODP	AP
a(H)	-0.14756	-0.18656	-0.01047	-0.18973	-0.01672	-0.12954	0.079188	1.67752	-3.37822	0.086429
a(Cl)	0.080051	-0.11963	0.152121	-12.3996	0.089494	0.428893	0.749693	3.423638	-11.9804	0.23872
a(Br)	0.183264	0.044889	0.572806	-27.3702	-0.21653	0.550528	1.136445	4.522534	-19.2469	****
a(F)	-0.06844	0.205501	0.197228	-6.47571	0.117109	-0.22263	0.620819	2.521654	-5.02663	0.296694
a(I)	0.30059	2.003523	0.30088	-42.9538	-0.31876	1.033486	****	1.781068	****	****
a(N)	-0.01828	-0.02395	0.139871	-4.71295	-0.35926	0.645083	0.09317	****	****	0.097458
a(O)	-0.08007	-0.21041	0.004962	-5.13567	-0.36447	0.224864	0.126631	0.424024	-3.62951	0.044628
a(P)	-0.20932	2.16253	1.181	-10.462	0.723115	-0.97773	****	****	****	****
a(S)	0.07225	0.247054	0.241384	-11.058	0.054859	1.524772	-0.24978	****	****	****
a(C)	0.285658	0.246449	0.052218	-4.42181	0.083797	0.325169	-0.03181	-3.05993	-1.59188	-0.10248
a(Si)	****	****	-0.02767	-11.0138	0.562317	1.052495	****	****	****	****
b	0.035109	0.253883	0.018103	-0.09321	0.283525	-0.42181	-0.10486	-0.01877	8.967731	-0.01884
c	0.184584	0.024195	-0.02677	-0.10952	-0.2022	0.241452	0.005087	-1.52848	0.22623	-0.12489
d	0	0	0	0	0.892444	2.998465	-0.25708	-0.52073	3.298083	-0.27126

Table B8 (continued). CI Method Based Property Models: Atom Contributions and Model Constants for the Properties÷ EUA_C, EUA_{NC}, ERA_C, ERA_{NC}, EFW_C, EFW_{NC}, ESW_C, ESW_{NC}, ENS_C, ENS_{NC}, EAS_C, and EAS_{NC}.

Parameter	EUA_C	EUA_{NC}	ERA_C	ERA_{NC}	EFW_C	EFW_{NC}	ESW_C	ESW_{NC}	ENS_C	ENS_{NC}	EAS_C	EAS_{NC}
a(H)	-0.0035	-0.0056	-0.0215	0.0073	-0.0337	-0.0046	0.0682	0.0429	-0.0527	-0.0333	-0.0407	-0.0142
a(Cl)	-0.3184	-0.3843	-0.5551	-0.5190	-0.4765	-0.3676	-0.3006	-0.5570	-0.3129	-0.3199	-0.3358	-0.2783
a(Br)	-0.0954	-0.2498	-0.6338	-0.3712	-0.4577	-0.0662	0.1593	-0.1070	-0.2447	-0.0237	-0.1894	-0.0239
a(F)	0.0843	-0.1457	-0.1594	-0.1494	-0.0821	-0.1904	-0.1000	-0.1074	-0.0716	-0.1681	-0.0844	-0.1567
a(I)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
a(N)	-0.1865	-0.0900	-0.3080	-0.2023	-0.3263	-0.0970	0.0363	0.0149	-0.2198	-0.1179	-0.2186	-0.0806
a(O)	-0.0263	-0.0875	-0.1808	-0.1550	0.0258	0.0332	0.2636	0.2549	-0.0090	-0.0565	0.0153	-0.0367
a(P)	0.6655	-0.8595	-0.0865	-0.7380	-0.2442	-1.5692	0.6766	-1.7886	0.2774	-1.0916	0.5614	-0.9972
a(S)	-0.1146	-0.0669	-0.3863	0.0192	-0.4479	-0.2211	0.1692	-0.0612	-0.2393	-0.1541	-0.2917	-0.1470
a(C)	-0.0373	-0.0205	-0.0748	-0.0580	-0.0651	-0.0401	-0.0824	-0.1454	0.1092	0.1268	0.0478	0.0751
a(Si)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
b	0.0605	0.0293	0.1630	0.0838	0.0670	-0.0763	-0.2888	-0.0771	-0.1154	-0.1819	-0.0141	-0.1722
c	0.0334	-0.0235	0.0567	-0.1104	0.1590	0.1340	0.2296	-0.0066	0.2944	0.2490	0.0721	0.1618
d	0	0	0	0	0	0	0	0	0	0	0	0

Appendix C

First-order, second-order and third-order group-contributions for the improved GC models for enthalpy of formation, enthalpy of fusion, and critical temperature.

Table C1. The total list of first-order groups and their contributions C_i for the improved GC-models for T_c, $\Delta_{fus}H$, and $\Delta_f H^o_{gas}$

	Group	T_{cIi}	$\Delta_{fus}H_{Ii}$	$\Delta_f H^o_{gasIi}$
	Universal constant	181.1926	4.5668	42.2361
1	CH_3	1.3098	-2.2062	-62.9440
2	CH_2	3.5123	2.5755	-20.6918
3	CH	3.9411	3.1165	16.5944
4	C	5.9549	5.7454	53.8782
5	CH_2=CH	4.9291	-2.8334	45.9239
6	CH=CH	7.4081	1.8399	81.5621
7	CH_2=C	6.7066	-0.4255	79.7497
8	CH=C	8.8436	2.4328	116.6983
9	C=C	11.5690	2.5307	159.8517
10	CH_2=C=CH	10.8679	3.6066	182.5487
11	CH_2=C=C	12.3243	7.7955	212.7318
12	C=C=C	****	****	****
13	CH≡C	6.6331	-3.6823	206.4532
14	C≡C	11.2790	6.5228	229.8654
15	aCH	3.9741	0.6272	6.7368
16	aC fused with aromatic ring	63.0089	1.8226	19.1902
17	aC fused with non-aromatic ring	34.2819	-2.8845	6.8644
18	aC except as above	11.1412	-4.2244	24.0669
19	aN in aromatic ring	8.6398	3.8679	63.4219
20	aC-CH_3	7.9636	2.3321	-28.7984
21	aC-CH_2	10.5924	4.1598	16.5000
22	aC-CH	12.9835	8.2531	50.1054
23	aC-C	7.2618	9.9143	90.7469
24	aC-CH=CH_2	15.1056	3.2469	70.4143
25	aC-CH=CH	20.7013	3.7263	111.4710
26	aC-C=CH_2	15.4591	4.3788	111.4917
27	aC-C≡CH	16.2674	1.7369	251.3800
28	aC-C≡C	****	****	251.3292
29	OH	11.0391	-0.1573	-192.6568
30	aC-OH	24.5442	8.8157	-174.1890
31	COOH	29.3694	4.4560	-411.8747
32	aC-COOH	44.9129	15.8129	-370.1686
33	CH_3CO	15.2685	3.4093	-192.2788
34	CH_2CO	15.0329	9.3701	-155.2908
35	CHCO	14.4146	9.8894	-113.0651
36	CCO	****	****	****
37	aC-CO	29.0076	4.8031	-104.6158
38	CHO	11.0543	7.5574	-144.9467
39	aC-CHO	25.2238	5.4909	-113.7486
40	CH_3COO	13.1652	4.4002	-401.8335
41	CH_2COO	17.6622	10.7091	-359.4218
42	CHCOO	14.6931	12.3357	-327.6094
43	CCOO	****	****	****
44	HCOO	11.3707	4.8125	-344.8540
45	aC-COO	20.1868	6.9998	-299.1613
46	aC-OOCH	****	****	****
47	aC-OOC	22.8276	8.9731	****
48	COO except as above	12.8523	11.4352	-300.7012
49	CH_3O	5.7385	1.5627	-173.6250
50	CH_2O	6.5266	5.0107	-147.9635
51	CH-O	6.1898	9.7594	-120.9651
52	C-O	8.3133	9.9692	-65.6020
53	aC-O	11.3665	2.6997	-86.7830
54	CH_2NH_2	11.7890	9.2490	-29.3809
55	$CHNH_2$	10.7482	6.0647	1.8859
56	CNH_2	11.1781	****	34.1501
57	CH_3NH	10.1603	0.6726	4.5407
58	CH_2NH	8.7273	8.0825	27.9307
59	CHNH	8.3325	5.2394	48.3349
60	CH_3N	7.6150	6.3855	52.4992
61	CH_2N	6.7610	7.8191	81.5377
62	aC-NH_2	29.1955	6.1648	7.6496
63	aC-NH	27.7448	4.0231	82.5978
64	aC-N	24.0929	8.1193	136.3567
65	NH_2 except as above	9.4690	1.5371	26.4749
66	CH=N	****	****	****
67	C=N	****	****	****

68	CH_2CN	20.9868	2.9966	73.6410
69	CHCN	19.7736	8.9295	108.8632
70	CCN	14.3190	7.8057	102.7072
71	aC-CN	29.0609	****	148.6680
72	CN except as above	15.0641	3.4523	149.8029
73	CH_2NCO	****	****	****
74	CHNCO	****	****	****
75	CNCO	****	****	****
76	aC-NCO	16.9793	****	-24.5200
77	$CH2NO_2$	24.3173	****	****
78	$CHNO_2$	23.6974	****	****
79	CNO_2	****	****	****
80	aC-NO_2	33.7161	6.5507	****
81	NO_2 except as above	24.3561	2.4755	****
82	ONO	****	****	****
83	ONO_2	****	5.5230	****
84	$HCON(CH_2)_2$	****	****	****
85	$HCONHCH_2$	****	****	****
86	$CONH_2$	65.3732	13.3493	-215.6939
87	$CONHCH_3$	51.2859	13.9393	-219.2922
88	$CONHCH_2$	****	****	-179.6447
89	$CON(CH_3)_2$	36.4594	8.0493	-206.3710
90	$CONCH_3CH_2$	****	****	****
91	$CON(CH_2)_2$	****	****	****
92	CONHCO	****	****	****
93	CONCO	****	****	****
94	aC-$CONH_2$	****	16.0569	-167.8646
95	aC-NH(CO)H	57.1017	7.8969	-131.1200
96	aC-N(CO)H	****	****	****
97	aC-CONH	****	****	****
98	aC-NHCO	****	****	****
99	aC-(N)CO	****	****	****
100	NHCONH	****	10.2900	****
101	NH_2CONH	****	11.1766	****
102	NH_2CON	****	11.4545	****
103	NHCON	****	16.3317	****
104	NCON	****	17.6579	3.9697
105	aC-$NHCONH_2$	****	15.9469	****
106	aC-NHCONH	****	28.0051	****
107	NHCO except as above	****	****	****
108	CH_2Cl	12.1137	1.0374	-90.5001
109	CHCl	11.2385	****	-59.0452
110	CCl	8.4668	3.4739	-30.7499
111	$CHCl_2$	16.4364	3.4019	-86.4791
112	CCl_2	****	2.1755	-86.4414
113	CCl_3	18.9340	-0.0107	-121.5922
114	CH_2F	6.8713	3.6393	-243.6922
115	CHF	5.8968	****	****
116	CF	****	****	****
117	CHF_2	6.2307	3.5093	-480.0922
118	CF_2	2.4389	****	****
119	CF_3	1.0125	-0.1516	-684.9514
120	CCl_2F	9.2917	-0.3751	-316.5747
121	HCClF	****	****	****
122	$CClF_2$	5.4272	-1.7318	-484.9514
123	aC-Cl	14.3744	4.1161	-22.9933
124	aC-F	2.9180	0.7641	-166.3727
125	aC-I	33.6475	2.0369	88.9800
126	aC-Br	20.7887	5.3956	32.8088
127	-I except as above	17.5140	2.4538	65.3679
128	-Br except as above	12.4108	0.4115	-22.4437
129	-F except as above	-1.2149	2.6466	-282.7819
130	-Cl except as above	5.2148	-0.1453	-62.5179
131	CHNOH	****	****	****
132	CNOH	****	****	****
133	aC-CHNOH	****	****	****
134	OCH_2CH_2OH	19.6216	4.2364	-360.3639
135	$OCHCH_2OH$	****	****	****
136	OCH_2CHOH	20.3130	11.5943	-301.9888
137	-O-OH	11.8153	****	-142.8582
138	CH_2SH	15.2432	6.1121	-25.7589
139	CHSH	14.5125	4.0447	10.4466
140	CSH	15.3644	3.8639	44.7201
141	aC-SH	24.9463	3.7469	35.5800
142	-SH except as above	12.0724	2.7371	0.7200
143	CH_3S	14.5342	4.0176	-19.0432
144	CH_2S	16.4898	4.8350	20.8004
145	CHS	****	****	****
146	CS	****	****	****
147	aC-S-	28.0177	****	****
148	SO	53.2680	13.7755	-66.8082
149	SO2	48.2220	18.1155	-301.7849
150	SO_3 (sulfite)	21.6630	****	-426.8647

151	SO_3 (Sulfonate)	****	****	****
152	SO_4 (Sulfate)	50.8511	****	-631.2647
153	aC-SO	****	****	****
154	aC-SO_2	****	12.7444	****
155	PH (phosphine)	****	****	****
156	P (Phospine)	****	****	****
157	PO_3 (Phospite)	****	****	****
158	PHO_3 (Phosponate)	****	****	****
159	PO_3 (Phosponate)	****	****	****
160	PHO_4 (Phospate)	****	****	****
161	PO_4 (Phospate)	49.1984	****	-1036.0289
162	aC-PO_4	****	24.0832	-948.7215
163	aC-P	178.5963	14.1632	128.7785
164	CO_3 (Carbonate)	16.6452	11.8455	-500.7024
165	C_2H_3O	13.1224	****	-140.2250
166	C_2H_2O	14.4014	****	****
167	C_2HO	****	****	****
168	CH_2 (cyclic)	5.1783	0.6680	-22.4553
169	CH (cyclic)	6.4465	2.4425	-70.2123
170	C (cyclic)	4.4952	-0.8189	89.7854
171	CH=CH (cyclic)	9.6905	0.6021	57.8276
172	CH=C (cyclic)	12.4869	6.9406	29.7633
173	C=C (cyclic)	****	****	****
174	CH_2=C (cyclic)	9.6127	-0.8209	****
175	NH (cyclic)	9.6586	5.2829	43.5426
176	N (cyclic)	10.6705	0.3032	92.2971
177	CH=N (cyclic)	29.9439	5.6191	52.7778
178	C=N (cyclic)	****	****	18.3043
179	O (cyclic)	6.3610	1.8958	-135.3785
180	CO (cyclic)	26.8756	4.6182	-190.6895
181	S (cyclic)	14.3371	2.8477	10.0093
182	SO_2 (cyclic)	71.1052	7.2460	-304.8175
183	>NH	****	****	****
184	-O-	11.0458	-0.4242	-130.2788
185	-S-	****	****	****
186	>CO	****	****	****
187	PO_2	****	****	****
188	CH-N	****	****	****
189	SiHO	****	****	****
190	SiO	****	****	****
191	SiH_2	****	****	****
192	SiH	****	****	****
193	Si	7.0378	****	****
194	$(CH_3)_3N$	****	****	****
195	N=N	****	****	194.7540
196	C_{cyclic}=N-	****	****	206.6854
197	C_{cyclic}=CH-	****	****	****
198	C_{cyclic}=NH	****	****	****
199	N=O	****	****	****
200	C_{cyclic}=C	17.6970	****	****
201	P=O	****	****	****
202	N=N	****	****	307.3707
203	C=NH	****	****	****
204	>C=S	****	****	111.1368
205	aC-CON	****	****	****
206	aC=O	****	-12.0627	****
207	aN-	****	****	****
208	-Na	****	****	****
209	-K	****	****	****
210	HCONH	****	****	****
211	CHOCH	****	****	****
212	C_2O	****	****	****
213	SiH_3	****	****	****
214	SiH_2O	****	****	****
215	CH=C=CH	****	****	****
216	CH=C=C	****	****	****
217	OP(=S)O	****	****	****
218	R	****	****	****
219	$CF_{2cyclic}$	****	****	****
220	CF_{cyclic}	****	****	****
221	CH3CNCHO	****	****	-80.6447

[a] The symbols T_{c1i}, $\Delta_f H^o_{gas1i}$, and $\Delta_{fus}H_{1i}$ represent the contributions (C_i) of the first-order groups for the corresponding properties

Table C2. The total list of second-order groups and their contributions D_j for the improved GC-models for T_c, $\Delta_{fus}H$, and $\Delta_f H^o_{gas}$

	Group	T_{c2j}	$\Delta_{fus}H_{2j}$	$\Delta_f H^o_{gas2j}$
1	$(CH_3)_2CH$	0.2131	1.2160	-2.6948
2	$(CH_3)_3C$	-0.6584	0.6678	-7.4243
3	$CH(CH_3)CH(CH_3)$	2.2187	-0.2243	6.0231
4	$CH(CH_3)C(CH_3)_2$	3.0440	-0.1584	9.0964
5	$C(CH_3)_2C(CH_3)_2$	5.3376	1.8110	17.7617
6	CH_n=CH_m-CH_p=CH_k (*k,m,n,p* in 0..2)	1.2092	4.4050	-24.5889
7	CH_3-CH_m=CH_n (*m,n* in 0..2)	0.3051	2.0443	-4.7405
8	CH_2-CH_m=CH_n (*m,n* in 0..2)	0.0163	0.5301	-4.6073
9	CH_p-CH_m=CH_n (*m,n* in 0..2; *p* in 0..1)	0.2147	0.3997	-3.8590
10	CHCHO or CCHO	-1.8154	-4.7266	-2.9650
11	CH_3COCH_2	-1.1561	0.6232	-5.7286
12	CH_3COCH or CH_3COC	-1.3928	2.5014	0.6758
13	CHCOOH or CCOOH	-7.0527	-3.9230	0.4212
14	CH_3COOCH or CH_3COOC	-1.4384	1.7615	-16.5891
15	CO-O-CO	-1.4135	-1.8764	-23.3372
16	CHOH	-2.3734	1.5893	-12.2882
17	COH	-5.1190	-0.3343	-21.6177
18	CH_3COCH_nOH (*n* in 0..2)	-2.9954	****	-2.6088
19	NCCHOH or NCCOH	2.6459	1.6672	40.6014
20	OH-CH_n-COO (*n* in 0..2)	****	****	****
21	$CH_m(OH)CH_n(OH)$ (*m,n* in 0..2)	1.9395	2.5878	-9.9789
22	$CH_m(OH)CH_n(NH_p)$ (*m,n,p* in 0..2)	4.7330	1.4337	-13.0193
23	$CH_m(NH_2)CH_n(NH_2)$ (*m,n* in 0..2)	2.8060	-0.4848	-3.0843
24	$CH_m(NH)CH_n(NH_2)$ (*m,n* in 1..2)	6.0273	-7.0114	1.7134
25	$H_2NCOCH_nCH_mCONH_2$ (*m,n* in 1..2)	****	****	****
26	$CH_m(NH_n)$-COOH (*m,n* in 0..2)	****	10.1282	10.2910
27	HOOC-CH_n-COOH (*n* in 1..2)	22.7603	9.4256	1.0050
28	HOOC-CH_n-CH_m-COOH (*n, m* in 1..2)	19.7185	34.4701	-0.0033
29	HO-CH_n-COOH (*n* in 1..2)	****	****	****
30	NH_2-CH_n-CH_m-COOH (*n, m* in 1..2)	****	****	****
31	CH_3-O-CH_n-COOH (n in 1..2)	6.6941	-2.6611	****
32	HS-CH-COOH	****	****	****
33	HS-CH_n-CH_m-COOH (*n, m* in 1..2)	7.7627	-0.7405	10.1492
34	NC-CH_n-CH_m-CN (*n, m* in 1..2)	28.1053	-6.8600	20.1819
35	OH-CH_n-CH_m-CN (*n, m* in 1..2)	9.5267	-2.4316	-0.8285
36	HS-CH_n-CH_m-SH (*n, m* in 1..2)	8.3396	-0.3911	-0.4183
37	COO-CH_n-CH_m-OOC (*n, m* in 1..2)	3.3863	****	-4.1856
38	OOC-CH_m-CH_m-COO (*n, m* in 1..2)	0.6682	0.1972	****
39	NC-CH_n-COO (*n* in 1..2)	6.5969	-0.4925	****
40	$COCH_nCOO$ (*n* in 1..2)	1.8441	****	-10.8630
41	CH_m-O-CH_n=CH_p (*m,n,p* in 0..3)	1.0218	3.7066	-18.4608
42	CH_m=CH_n-F (*m,n* in 0..2)	2.3910	****	55.7118
43	CH_m=CH_n-Br (*m,n* in 0..2)	-3.7342	3.3551	12.6537
44	CH_m=CH_n-I (*m,n* in 0..2)	****	****	****
45	CH_m=CH_n-Cl (*m,n* in 0..2)	0.3059	1.5121	-8.3533
46	CH_m=CH_n-CN (*m,n* in 0..2)	-0.5497	****	-43.8706
47	CH_n=CH_m-COO-CH_p (*m,n,p* in 0..3)	-1.6624	****	-50.0861
48	CH_m=CH_n-CHO (*m,n* in 0..2)	2.5291	****	****
49	CH_m=CH_n-COOH (*m,n* in 0..2)	0.8098	6.7546	-11.4295
50	aC-CH_n-X (*n* in 1..2) X: Halogen	8.4030	1.4680	-11.2021
51	aC-CH_n-NH_m (*n* in 1..2; *m* in 0..2))	3.5449	****	-2.9485
52	aC-CH_n-O- (*n* in 1..2)	0.3129	0.7326	3.4875
53	aC-CH_n-OH (*n* in 1..2)	-3.2132	0.7823	-4.6439
54	aC-CH_n-CN (*n* in 1..2)	11.2936	****	-26.1229
55	aC-CH_n-CHO (*n* in 1..2)	****	****	****
56	aC-CH_n-SH (*n* in 1..2)	8.3472	****	****
57	aC-CH_n-COOH (*n* in 1..2)	8.7164	3.3173	****
58	aC-CH_n-CO- (*n* in 1..2)	****	****	10.9474
59	aC-CH_n-S- (*n* in 1..2)	****	****	****
60	aC-CH_n-OOC-H (*n* in 1..2)	5.2654	****	3.1340
61	aC-CH_m-NO_2 (*n* in 1..2)	****	****	****
62	aC-CH_n-$CONH_2$ (*n* in 1..2)	****	****	****
63	aC-CH_n-OOC (*n* in 1..2)	3.7316	****	0.2535
64	aC-CH_n-COO (*n* in 1..2)	****	****	****
65	aC-SO_2-OH	****	****	-362.2632

66	aC-CH(CH_3)$_2$	-1.1750	-0.3867	1.1849
67	aC-C(CH_3)$_3$	6.9137	-1.3538	-0.6319
68	aC-CF_3	-5.5827	****	****
69	(CH_n=C)$_{cyc}$-CHO (*n* in 0..2)	9.2509	-3.8717	-0.5019
70	(CH_n=C)$_{cyc}$-COO-CH_m (*n,m* in 0..3)	****	****	****
71	(CH_n=C)$_{cyc}$-CO- (*n* in 0..2)	****	****	****
72	(CH_n=C)$_{cyc}$-CH_3 (*n* in 0..2)	-0.8979	****	11.1992
73	(CH_n=C)$_{cyc}$-CH_2 (*n* in 0..2)	-1.9146	-0.0025	****
74	(CH_n=C)$_{cyc}$-CN (*n* in 0..2)	****	****	****
75	(CH_n=C)$_{cyc}$-Cl (*n* in 0..2)	****	****	****
76	CH_{cyc}-CH_3	-0.7034	1.3677	76.7315
77	CH_{cyc}-CH_2	0.2840	0.7024	71.6246
78	CH_{cyc}-CH	0.8083	3.7351	75.7857
79	CH_{cyc}-C	2.9225	****	****
80	CH_{cyc}-CH=CH_n (*n* in 1..2)	-1.3813	****	73.4788
81	CH_{cyc}-C=CH_n (*n* in 1..2)	****	****	****
82	CH_{cyc}-Cl	-4.2213	****	****
83	CH_{cyc}-F	****	****	****
84	CH_{cyc}-OH	0.7081	-4.9174	63.4977
85	CH_{cyc}-NH_2	-4.0695	2.7331	25.6660
86	CH_{cyc}-NH-CH_n (*n* in 0..3)	-3.5875	3.8584	49.5002
87	CH_{cyc}-N-CH_n (*n* in 0..3)	****	****	****
88	CH_{cyc}-SH	2.5779	0.1939	60.3009
89	CH_{cyc}-CN	****	****	****
90	CH_{cyc}-COOH	****	****	****
91	CH_{cyc}-CO	****	****	****
92	CH_{cyc}-NO_2	****	****	****
93	CH_{cyc}-S-	****	****	****
94	CH_{cyc}-CHO	****	****	****
95	CH_{cyc}-O-	****	****	****
96	CH_{cyc}-OOCH	****	****	****
97	CH_{cyc}-COO	****	****	****
98	CH_{cyc}-OOC	-1.8146	****	55.8744
99	C_{cyc}-CH_3	0.3557	1.2470	-23.5741
100	C_{cyc}-CH_2	1.5888	****	-27.4799
101	C_{cyc}-OH	-5.6809	4.3689	-55.7819
102	>N_{cyc}-CH_3	3.3871	****	-0.5124
103	>N_{cyc}-CH_2	****	****	-5.9388
104	AROMRINGs^{1}s^2	-0.8547	-2.4702	5.7727
105	AROMRINGs^{1}s^3	-1.3540	-3.5292	3.8909
106	AROMRINGs^{1}s^4	-0.6007	-0.9685	3.7393
107	AROMRINGs^{1}s^{2}s^3	0.8117	-2.1657	12.7833
108	AROMRINGs^{1}s^{2}s^4	-0.3754	-2.4122	9.5412
109	AROMRINGs^{1}s^{3}s^5	-1.9551	-3.3965	7.3474
110	AROMRINGs^{1}s^{2}s^{3}s^4	6.0144	-3.9197	24.1771
111	AROMRINGs^{1}s^{2}s^{3}s^5	2.6076	-4.4297	18.9321
112	AROMRINGs^{1}s^{2}s^{4}s^5	1.7160	5.8503	16.4521
113	PYRIDINEs^2	-1.7068	-3.5558	-4.9041
114	PYRIDINEs^3	2.6544	0.9042	0.5481
115	PYRIDINEs^4	2.8782	0.6178	1.1594
116	PYRIDINEs^{2}s^3	0.7622	****	4.2833
117	PYRIDINEs^{2}s^4	-0.9449	****	-0.2824
118	PYRIDINEs^{2}s^5	-1.4707	****	-0.5991
119	PYRIDINEs^{2}s^6	-5.2254	-4.9406	-8.8798
120	PYRIDINEs^{3}s^4	7.0354	****	0.4206
121	PYRIDINEs^{3}s^5	3.3568	****	2.9436
122	PYRIDINEs^{2}s^{3}s^6	****	****	****
123	(CH^n=CH^m)^{cyc}-COOH	****	****	****
124	AROMRINGs^{1}s^{2}s^{3}s^{4}s^5	****	-6.1845	****
125	aC-NHCOCH_2N	****	****	****
126	(N=C)$_{cyc}$-CH_3	****	****	45.7854
127	aC-CONH(CH_2)$_2$N	****	****	****
128	aC-SO_2NH_n (*n*>=0;*n*<3)	****	****	****
129	aC-SO_2NH_n (*n*>=0;*n*<3)	****	****	****
130	aC-SO_2NH_n (*n*>=0;*n*<3)	****	****	****

[a] The symbols T_{c2j}, $\Delta_f H^o_{gas2j}$, and $\Delta_{fus}H_{2j}$ represent the contributions (D_j) of the second-order groups for the corresponding properties

Table C3. The total list of third-order groups and their contributions E_k for the improved GC-models for T_c, $\Delta_{fus}H$, and $\Delta_f H^o_{gas}$

	Group	T_{c3k}	$\Delta_{fus}H_{3k}$	$\Delta_f H^o_{gas3k}$
1	HOOC-$(CH_n)_m$-COOH (m>2, n in 0..2)	16.8529	-2.2309	0.0803
2	NH_n-$(CH_n)_m$-COOH (m>2, n in 0..2)	****	****	****
3	NH_2-$(CH_n)_m$-OH (m>2, n in 0..2)	6.0863	-2.9096	0.1851
4	OH-$(CH_n)_m$-OH (m>2, n in 0..2)	2.5570	4.2070	0.8697
5	OH-$(CH_p)_k$-O-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****
6	OH-$(CH_p)_k$-S-$(CH_n)_m$-OH (m,k>0; p,n in 0..2)	****	****	****
7	OH-$(CH_p)_k$-NH_x-$(CH_n)_m$-OH (m,k>0; p,n,x in 0..2)	****	****	****
8	CH_p-O-$(CH_n)_m$-OH (m>2; n,p in 0..2)	****	****	****
9	NH_2-$(CH_n)_m$-NH_2 (m>2; n in 0..2))	0.5612	7.0131	0.4928
10	NH_k-$(CH_n)_m$-NH_2 (m>2; k in 0..1; n in 0..2)	****	****	****
11	SH-$(CH_n)_m$-SH (m>2; n in 0..2)	****	****	****
12	NC-$(CH_n)_m$-CN (m>2)	27.4956	3.3667	1.2696
13	COO-$(CH_n)_m$-OOC (m>2; n in 0..2)	****	****	****
14	aC-$(CH_n$=$CH_m)_{cyc}$ (fused rings) (n,m in 0..1)	-0.4615	1.6879	16.5161
15	aC-aC (different rings)	15.5193	9.4463	18.8705
16	aC-CH_{ncyc} (different rings) (n in 0..1)	5.3014	9.2891	82.5902
17	aC-CH_{ncyc} (fused rings) (n in 0..1)	2.1514	5.8071	20.4781
18	aC-$(CH_n)_m$-aC (different rings) (m>1; n in 0..2)	13.1293	11.3811	0.3962
19	aC-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****
20	CH_{cyc}-CH_{cyc} (different rings)	6.4932	****	****
21	CH_{cyc}-$(CH_n)_m$-CH_{cyc} (different rings) (m>0; n in 0..2)	****	****	****
22	$CH_{multiring}$	3.1628	-1.1783	71.0836
23	$C_{multiring}$	****	****	****
24	aC-CH_m-aC (different rings) (m in 0..2)	7.6591	2.7334	9.6761
25	aC-$(CH_m$=$CH_n)$-aC (different rings) (m,n in 0..2)	12.0246	17.3488	-3.2418
26	$(CH_m$=C$)_{cyc}$-CH=CH-$(C$=$CH_n)_{cyc}$ (different rings)	****	****	****
27	$(CH_m$=C$)_{cyc}$-CH_p-$(C$=$CH_n)_{cyc}$ (different rings)	****	****	****
28	aC-CO-aC (different rings)	5.4806	5.4820	****
29	aC-CH_m-CO-aC (different rings) (m in 0..2)	4.9533	****	29.9450
30	aC-CO-$(C$=$CH_n)_{cyc}$ (different rings) (n in 0..1)	****	****	****
31	aC-CO-CO-aC (different rings)	****	2.2445	****
32	aC-CO_{cyc} (fused rings)	****	9.7312	15.3954
33	aC-CO-$(CH_n)_m$-CO-aC (different rings) (m>0; n in 0..2)	****	****	11.4694
34	aC-CO-CH_{ncyc} (different rings) (n in 0..1)	****	****	****
35	aC-CO-NH_n-aC (different rings) (n in 0..1)	****	22.9951	****
36	aC-NH_nCONH_m-aC (different rings) (n,m in 0..1)	****	****	****
37	aC-CO-N_{cyc} (different rings)	****	****	****
38	aC-S_{cyc} (fused rings)	17.9458	5.9537	9.2049
39	aC-S-aC (different rings)	****	****	****
40	aC-PO_n-aC (different rings) (n in 0..4)	****	****	****
41	aC-SO_n-aC (different rings) (n in 1..4)	****	****	****
42	aC-NH_{ncyc} (fused rings) (n in 0..1)	3.5500	6.6580	2.8160
43	aC-NH-aC (different rings)	13.8042	8.0620	7.0683
44	aC-$(C$=$N)_{cyc}$ (different rings)	****	****	****
45	aC-$(N$=$CH_n)_{cyc}$ (fused rings) (n in 0..1)	****	-1.0603	-10.2580
46	aC-$(CH_n$=$N)_{cyc}$ (fused rings) (n in 0..1)	****	****	51.9557
47	aC-O-CH_n-aC (different rings) (n in 0..2)	****	****	****
48	aC-O-aC (different rings)	6.5934	-1.4087	5.1123
49	aC-CH_n-O-CH_m-aC (different rings) (n,m in 0..2)	0.2417	****	****
50	aC-O_{cyc} (fused rings)	8.6032	11.3607	19.0134
51	AROM.FUSED[2]	-50.1114	0.9489	10.5077
52	AROM.FUSED[2]s^1	-40.7738	-0.0595	-1.8447
53	AROM.FUSED[2]s^2	-43.8246	5.9021	6.8914
54	AROM.FUSED[2]s^2s^3	-7.4882	6.7448	****
55	AROM.FUSED[2]s^1s^4	****	-0.3686	****
56	AROM.FUSED[2]s^1s^2	****	****	****
57	AROM.FUSED[2]s^1s^3	-5.3001	****	****
58	AROM.FUSED[3]	-69.3585	1.0817	12.8155
59	AROM.FUSED[4a]	-67.7173	9.3426	56.9458

60	AROM.FUSED[4a]s^1	****	****	53.4281
61	AROM.FUSED[4a]s^1s^4	****	****	46.8901
62	AROM.FUSED[4p]	-57.2711	-2.6322	-13.2545
63	AROM.FUSED[4p]s^3s^4	****	****	****
64	PYRIDINE.FUSED[2]	0.7998	-6.1033	2.6816
65	PYRIDINE.FUSED[2-iso]	14.8858	-4.2395	13.5421
66	PYRIDINE.FUSED[4]	****	-2.5881	35.2804
67	aC-N-CH_{cyc} (different rings)	****	****	****
68	N multiring	****	****	45.3816
69	N_{cyc}-$(CH_2)_3$-N_{cyc} (different rings)	****	****	****
70	aC-$COCH_2CH_2$-aC (different rings)	****	****	****
71	aC-O-$(CH_2)_2$-N_{cyc} (different rings)	****	****	****
72	aC-CH(OH)$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****
73	Ncyc-$(CH_2)_2$-CH_{cyc} (different rings)	****	****	****
74	aC-$CONHCH_2$-CH_{cyc} (different rings)	****	****	****
75	CH_3 – $(CH_2)_m$ –I m in 0,1,2…m	0.0985	****	-34.6802
76	aC– (N_{cyc} =CH_{cyc}) –S_{cyc}	****	0.4551	-24.4102
77	aC– (N_{cyc} =CH_{cyc}) –O_{cyc}	****	****	25.5675
78	CH_3– $(CH_2)_m$– (N=N)m in 0,1,2…m	****	****	-75.0189
79	aC– (N=N)–aC	****	****	-14.5083
80	aN– aC –CN	****	3.1669	38.8105
81	CH_3 – $(CH_2)_m$ –COO–aC	0.4157	****	14.4053
82	CH_3O –$(CH_2)_m$ –O CH_2 CH_2OH	-0.0233	2.1830	-25.8636
83	$(aN–aCH)_3$	****	****	-26.8423
84	aC– aN –aC	****	****	-0.1817
85	aC–$CONH_2$–aC–aN	****	****	-1.5721
86	CH_3 –CH_2COO m in 0,1,2…m	-2.4668	****	-13.3623
87	aC–aN–aC–$(CH_3)_m$ -$(OH)_m$ –$(Cl)_m$ m in 0,1	-21.5516	****	-6.8457
88	OH–$(aC)_m$–aN m in 1,2	****	****	2.7171
89	aC–N_{cyc} –(CH_{cyc}=CH_{cyc})	****	****	-14.5362
90	Pyrazine in fused rings	-21.5516	****	6.9165
91	CH_3 – $(CH_2)_m$ – CH_2COO m in 1,2…m	-4.0172	-1.0343	10.7110
92	CH_3 – $(CH_2)_m$ –aC–NH m in 0,1,2…m	-5.9026	-2.4755	-7.9546
93	aC– (aC– CH_3) in fused rings	-2.6304	-4.9833	7.5283
94	$(aC\text{-}Cl)_2$– aCH– aC-Cl	-22.8633	****	19.7547
95	CH_3 – $(CH_2)_m$ – CH_2O m in 1,2…m	-0.4922	-1.0343	-4.5094
96	Cl–CH-Cl	0.5119	****	-21.2742
97	NH_{cyc} –$(CH_{cyc})_m$ m in 3,4…m	1.5020	****	-1.5119
98	CH_3–N_{cyc}–CO_{cyc}	4.3872	****	-13.7445
99	aC-NH_{cyc} in different rings	****	****	6.4066
100	5 member ring	-8.4548	-3.2910	0.5335
101	6 member ring	-7.9472	-3.2805	-16.7881
102	7 member ring	-8.4597	-3.2510	-14.6925
103	$(CH_3)_3C$– $(CH_2)_m$–$(CH_3)_3C$ m in 1,2…m	0.9814	-0.3447	14.9167
104	$(CH_3)_3C$–CH=C	0.2357	-0.0804	17.5714
105	(C_n H_{2n})n in 4,5,6	-8.6319	-4.7993	-10.6186
106	(C_n H_{2n}) n in 7,8,9	-7.7785	-7.5177	16.5373
107	(C_n H_{2n}) n in 10, 11,….n	****	****	4.8331
108	(CH_{cyc})-CO_{cyc}-(CH_{cyc})	****	****	45.9867

[a] The symbols T_{c3k}, $\Delta_f H^o_{gas3k}$, and $\Delta_{fus}H_{3k}$ represent the contributions (E_k) of the third-order groups for the corresponding properties

Appendix D

Table D1. List of molecular-group compounds for which developed GC model for $\Delta_f H^o_{gas}$ is not applicable

Sl. No.	Molecular-group
1	Water
2	Furfural
3	Formaldehyde
4	Dichloromethylene
5	Difluoromethylene
6	Methanethiol
7	Thioformaldehyde
8	Carbon disulfide
9	Formic acid
10	Thiophene dimethylsulfoxide
11	Urea
12	methylamine
13	Acetonitrile
14	Methane, dichloro-
15	Methane, trifluoro-
16	trichlorofluoromethane
17	nitromethane
18	carbon disulfide
19	Ethylene oxide
20	Methanol
21	Pyridine
22	Carbon tetrachloride
23	Morpholine
24	Dimethylformamide
25	ETHYLENE GLYCOL
26	Chloroform
27	dimethyl sulfide
28	Thiophene
29	dimethyl sulfoxide
30	N-methylpyrolidone

Appendix E

Mathematical model for short-path evaporator developed by Sales-Cruz and Gani (2006)

In the following section, the steady-state process model developed by Sales-Cruz and Gani (2006) is discussed. The momentum, energy and mass balances are applied to the evaporation and condensation films to derive the mathematical model for short-path evaporation process (see Figure E.1).

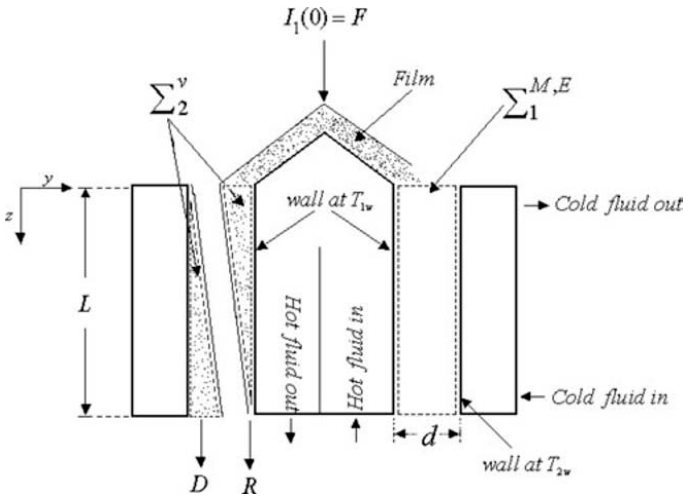


Figure E.1 Balance volumes for the short-path evaporator. The balance volume 1 is for the mass and energy balance; the balance volume 2 is for the momentum balance (Adapted from Sales-Cruz and Gani, 2006).

Momentum balance:

In most cases of short-path evaporation, the evaporating liquid is highly viscous and hence the corresponding Reynolds numbers are small. The Navier–Stokes equation at steady state for laminar flow regime describing the velocity profile of falling film is,

$$v(y,z)\frac{\partial^2 v(y,z)}{\partial y^2} = -g \qquad \text{(E.1)}$$

where y and z are the radial and axial coordinates respectively, v is the velocity and g is the gravitational constant.

Rate of evaporation:

The rate of evaporation is obtained from the continuity equation in terms of flow rate (I_i) for each component i,

$$\frac{\partial I_i(z)}{\partial z} = -2\pi R k_i \qquad \text{Where } i = 1, 2, 3\ldots\ldots N \tag{E.2}$$

Where, the effective rate of evaporation of each component (k_i) is calculated through a modified Langmuir – Knudsen equation (Kawala and Stephan, 1989).

$$k_i = \frac{\gamma_i P_{vap,i}}{\sqrt{2\pi R_g M W_i T_s(z)}} \left(\frac{P}{P_{ref}}\right) \left\{1-(1-F)\left[1-e^{d/k\beta}\right]^n\right) \tag{E.3}$$

Equation (E.3) contains a factor (P/P_{ref}) for correcting the vacuum pressure, as well as a correction that takes into account the anisotropic properties of the vapor. Here, β is the mean path of vapor molecule, d is the distillation gap, n is the number of intermolecular collision, F is the surface ratio, and k is the degree of anisotropy of the vapor phase in the space between the evaporator and the condenser. The effective rate of evaporation given by Eq. (E.3) also depends on some mixture properties (activity coefficient γ, vapor pressure P_{vap}, and molecular weight MW of each pure component) as well as on design parameters (the radius of the evaporator inside cylinder R and the surface temperature T_s).

Energy balance:

The temperature (T) profile in the falling film is given by the equation,

$$v(y,z)\frac{\partial T(y,z)}{\partial z} = \frac{\lambda}{\rho Cp}\left[\frac{\partial^2 T(y,z)}{\partial y^2} + \frac{\partial^2 T(y,z)}{\partial z^2}\right] \tag{E.4}$$

With following boundary conditions: At $z = 0$, the temperature corresponds to the liquid feed temperature and at $y = 0$, the temperature corresponds to the evaporator wall temperature.

Mass balance:

The composition (C_i) profiles for each component are calculated from the diffusion equation,

$$v(y,z)\frac{\partial C_i(y,z)}{\partial z} = D_i\left[\frac{\partial^2 C_i(y,z)}{\partial y^2} + \frac{\partial^2 C_i(y,z)}{\partial z^2}\right] \qquad \text{Where } i = 1, 2, 3\ldots\ldots N \tag{E.5}$$

Where D_i is the diffusion coefficient for the i^{th} component, and N is the total number of pure components.

Appendix F

This appendix contains a list of conference presentations and publications related to this PhD project. The results from this PhD work disseminated in the form of 4 research articles in scientific journals, and 3 articles published as conference proceedings are listed in section F1. Furthermore, the conference presentation (both oral and poster) presented in various international conferences during the course of PhD study are listed under section F2.

F1. Journal publications/ peer reviewed conference proceedings:

1. **Hukkerikar, A. S**., Sarup, B., Ten Kate, A., Abildskov, J., Sin, G., and Gani, R. Group contribution^{+} (GC^{+}) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilibria*. 2012, 321, 25-43.
2. **Hukkerikar, A. S**., Kalakul, S., Sarup, B., Young, D. M., Sin, G., and Gani, R. Estimation of environment-related properties of chemicals for design of sustainable processes: Development of group-contribution^{+} (GC^{+}) models and uncertainty analysis. *Journal of Chemical Information and Modeling.* 2012, 52 (11), 2823-2839.
3. **Hukkerikar, A. S**., Meier, R. J., Sin, G., and Gani, R. A method to estimate the enthalpy of formation of organic compounds with *chemical accuracy*, *Fluid Phase Equilibria*. 2013, 348, 23-32.
4. Cunico, L. P., **Hukkerikar, A. S**., Ceriani, R., Sarup, B., and Gani, R. Molecular structure-based methods of property prediction in application to lipids: A review and refinement. *Fluid Phase Equilibria*. 2013 (In Press).
5. Diaz-Tovar, C., Mustaffa, A., **Hukkerikar, A**.**S**, Quaglia, A., Sin, G., Kontogeorgis, G, Sarup, B. Gani, R. Lipid Processing Technology: Building a Multilevel Modeling Network. *Computer Aided Chemical Engineering*. 2011, 29, 256-260.
6. **Hukkerikar, A. S**., Jones, M., Sarup, B., Abildskov, J., Sin, G., and Gani, R. Sensitivity of process design due to uncertainties in property estimates. *Computer Aided Chemical Engineering*. 2012, 31, 200-204.
7. **Hukkerikar, A. S**., Sarup, B., Sin, G., and Gani, R. Molecular structure based property modeling: Development/improvement of property models through a systematic property-data-model analysis. 2013, PPEPPD-2013 conference.

F.2 Conference contributions (Oral and Poster presentations)

1. Diaz-Tovar, C., Mustaffa, A., **Hukkerikar, A**.**S**, Quaglia, A., Sin, G., Kontogeorgis, G, Sarup, B. Gani, R. “Lipid Processing Technology: Building a Multilevel Modeling

Network", Type: Oral, Presented by: Diaz-Tovar, C., Presented at: ESCAPE-21 conference, Chalkidiki, Greece.

2. **Hukkerikar, A. S**., Sarup, B., Sin, G., Gani, R. "Estimation of properties of pure components using improved group-contribution based and atom connectivity index based property models and uncertainty analysis", Type: Oral, Presented at: 25th European Symposium on Applied Thermodynamics (ESAT), Saint Petersburg, Russia.

3. **Hukkerikar, A. S**., Sarup, B., Sin, G., Gani, R. "A systematic methodology for uncertainty analysis of group contribution based and atom connectivity index based models for estimation of properties of pure components", Type: Oral, Presented at: 8th European Congress of Chemical Engineering (ECCE), Berlin, Germany.

4. Mustaffa, A., Diaz-Tovar, C., **Hukkerikar, A.S**, Quaglia, A., Sin, G., Kontogeorgis, G, Sarup, B. Gani, R. "Building a multilevel modeling network for lipid processing systems", Type: Oral, Presented by: Mustaffa, A., Presented at: 4th International Conference on Modeling, Simulation and Applied Optimization, Kuala Lumpur.

5. **Hukkerikar, A. S**., Sarup, B., J. Abildskov, Sin, G., Gani, R. "Development of property models with uncertainty estimate for process design under uncertainty", Type: Oral, Presented at: 2011 AIChE annual meeting, Minneapolis, USA.

6. **Hukkerikar, A. S**., Sarup, B., J. Abildskov, Sin, G., Gani, R. "Development of property models with uncertainty estimate for reliable product-process design", Type: Oral, Presented at: Industrial Use of Molecular Thermodynamics (InMoTher 2012), Lyon, France.

7. **Hukkerikar, A. S**., Sarup, B., J. Abildskov, Sin, G., Gani, R. "Estimation of properties of pure components using improved group-contribution^{+} (GC^{+}) based models and uncertainty analysis", Type: Oral, Presented at: 18th Symposium of Thermophysical Properties held in Colorado, USA

8. **Hukkerikar, A. S**., Sarup, B., J. Abildskov, Sin, G., Gani, R. "Effect of uncertainties in physical property estimates on process design - sensitivity analysis", Type: Oral, Presented at: 18th Symposium of Thermophysical Properties held in Colorado, USA

9. **Hukkerikar, A. S**., Jones, M, Sarup, B., J. Abildskov, Sin, G., Gani, R. "Sensitivity of process design due to uncertainties in property estimates" Type: Oral and poster, Presented by: Sin, G., Presented at: PSE-2012, Singapore.

10. Jones, M, **Hukkerikar, A. S**., Sin, G., Gani, R. “Sensitivity of process design to uncertainties in property models applied to extractive distillation”, Type: Oral, Presented by: Jones, M., Presented at: CHISA-2012, Prague.

11. **Hukkerikar, A. S**., Sarup, B., Young, D., Sin, G., Gani, R. “Group contribution^{+} (GC^{+}) based estimation of environment-related properties for design of sustainable processes: Development of property models and uncertainty analysis”, Type: Oral, Presented at: 2012 AIChE annual meeting, Pittsburgh, USA.

12. Cunico, L., **Hukkerikar, A. S**., Sin, G., Gani, R."Molecular structure based physical properties modelling", Presented by: Gani, R., Type: Oral, Presented at: MTMS'12, Higashi-Hiroshima.

13. **Hukkerikar, A. S**., Sarup, B., Sin, G., Gani, R. “Molecular structure based property modeling: Development/improvement of property models through a systematic property-data-model analysis”, Type: Poster, Presented at: PPEPPD-2013, Iguazu, Argentina.

This PhD-project was carried out at CAPEC, the Computer Aided Product-Process Engineering Center. CAPEC is committed to research, to work in close collaboration with industry and to participate in educational activities. The research objectives of CAPEC are to develop computer-aided systems for product/process simulation, design, analysis and control/operation for chemical, petrochemical, pharmaceutical and biochemical industries. The dissemination of the research results of CAPEC is carried out in terms of computational tools, technology and application. Under computational tools, CAPEC is involved with mathematical models, numerical solvers, process/operation mathematical models, numerical solvers, process simulators, process/product synthesis/design toolbox, control toolbox, databases and many more. Under technology, CAPEC is involved with development of methodologies for synthesis/design of processes and products, analysis, control and operation of processes, strategies for modelling and simulation, solvent and chemical selection and design, pollution prevention and many more. Under application, CAPEC is actively involved with developing industrial case studies, tutorial case studies for education and training, technology transfer studies together with industrial companies, consulting and many more.

Further information about CAPEC can be found at www.capec.kt.dtu.dk.

Computer Aided Process Engineering Center
Department of Chemical and Biochemical Engineering
Technical University of Denmark
Søltofts Plads, Building 229
DK-2800 Kgs. Lyngby
Denmark

Phone: +45 4525 2800
Fax: +45 4525 4588
Web: www.capec.kt.dtu.dk

ISBN: 978-87-93054-11-0